Exhibit 1

UNITED STATES DISTRICT COURT FOR THE WESTERN DISTRICT OF NEW YORK

BLACK LOVE RESISTS IN THE RUST, *et al.*, individually and on behalf of a class of all others similarly situated,

Plaintiffs

v.

CITY OF BUFFALO, N.Y., et al.,

Defendants.

No. 1:18-cv-00719-CCR

EXPERT REPORT OF DAVID BJERK, PH.D.

TABLE OF CONTENTS

I. QUALIFICATIONS AND BACKGROUND

1. I am the Russell S. Bock Professor of Public Economics and Taxation at Claremont McKenna College, Chair of the Robert Day School of Economics and Finance at Claremont McKenna College, a Fellow of the Society for Empirical Legal Studies, and a Research Fellow at the Institute for Study of Labor.

2. I have a Ph.D. and M.S. in Economics from the University Wisconsin-Madison, and a B.A. in Economics from Carleton College.

3. Claremont McKenna College is a private liberal arts school, historically ranking as one of the top liberal arts schools in the country. Claremont McKenna is known to have a focus on economics, with over half the students majoring in economics. It has the largest economics department among all liberal arts schools in the United States.

4. Prior to coming to Claremont McKenna, I was a Research Fellow in Population Studies at the RAND Corporation in Santa Monica, California, and an Assistant Professor of Economics at McMaster University in Hamilton, Ontario. Prior to attending graduate school, I worked as an Assistant Economist at the New York City office of Skadden, Arps, Meagher, and Flom LLP and as a Research Assistant at Lexecon Inc. in Chicago, Illinois.

5. Over the course of my career, I have taught courses in the Economics of Poverty, Inequality, and Discrimination; the Economics of Crime and Criminal Justice; Public Economics; Labor Economics; Intermediate Microeconomic Theory; Research Methods; and the Economics seminar and tutorial in the Politics, Philosophy, and Economics (PPE) program.

6. I am currently an Associate Editor at the *International Review of Law and Economics*. I have reviewed well over one hundred and fifty papers for various economics, law, and criminology journals, as well as numerous research proposals for a variety of foundations including the National Science Foundation, the Social Science and Humanities Research Council (the main academic research foundation in Canada), and the Israeli Science Foundation. I was named a Distinguished Referee by the *Journal of Legal Studies* for 2016-2018.

7. I have published extensively in internationally known, peer-reviewed academic journals in economics, legal studies, and criminology, including in the *Economic Journal*, the *Journal of Law and Economics*, the *Journal of Public Economics*, the *Journal of Legal Studies*, the *Journal of Human Resources*, and the *Journal of Quantitative Criminology*. According to Google Scholar, my work has been cited by over 1,380 other studies.

8. My research focuses on crime and the criminal justice system, the relationship between poverty and crime, and racial inequality and discrimination in different contexts including the labor market and policing. My research uses mathematical modeling, data analysis, and statistical inference to analyze these issues.

9. I have published several papers analyzing racial discrimination, including "The Differing Nature of Black-White Wage Inequality" (*Journal of Human Resources*), "Racial Profiling, Statistical Discrimination, and the Effect of a Colorblind Policy on the Crime Rate" (*Journal of Public Economic Theory*), "Glass Ceilings or Sticky Floors? Statistical Discrimination in a Dynamic Model of Promotion and Hiring" (*Economic Journal*), and "What Can DNA Exonerations Tell Us About Racial Differences in Wrongful Conviction Rates?" (*Journal of Law and Economics*). I have also published papers specifically discussing statistical methodologies in the context of discrimination and criminology, including "Using Samples-of-Opportunity to Assess Gender Bias in Principal Evaluations of Teachers: A Cautionary Tale" (*Journal of Labor Research*), "How Much Can We Trust Causal Interpretations of Fixed Effects Estimators in the Context of Criminality" (*Journal of Quantitative Criminology*), and "Measuring the Relationship Between Youth Criminal Participation and Household Economic Resources" (*Journal of Quantitative Criminology*). Other statistical analyses I have published in the areas of crime, policing, plea bargaining, and sentencing include "The Market for Mules: Risk and Compensation of Cross Border Drug Mules" (*International Law and Economics Review*), "Does Greater Police Funding Help Catch More Murderers?" (*Journal of Empirical Legal Studies*), "Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion Under Mandatory Minimum Sentencing" (*Journal of Law and Economics*), and "Mandatory Minimums and the Sentencing of Federal Drug Crimes" (*Journal of Legal Studies*)

10. My full *curriculum vitae* are presented in Exhibit A, which identifies all of my papers published within the last 10 years. I have not previously testified in litigation. I am currently being compensated at a rate of $250 per hour. Before January 1, 2023, I was compensated at a rate of $175 per hour. My compensation is in no way dependent on the conclusions that I reach or the outcome of this case.

II. ISSUES ADDRESSED

11. The plaintiffs in this case are current and former residents of the City of Buffalo who allege that the Buffalo Police Department (hereafter BPD) has engaged in certain unlawful

policies and practices relating to the enforcement of traffic and motor-vehicle-safety laws. These alleged policies and practices include (i) issuing citations for the purposes of generating revenue, rather than for legitimate law-enforcement purposes; (ii) discriminating on the basis of race when conducting stops and issuing citations; and (iii) discriminating on the basis of race in deciding when and where to locate traffic checkpoints that motorists must pass through and could lead to stops and citations or searches.

12. The plaintiffs claim that these alleged policies and practices disproportionately impacted Black and Hispanic residents (hereafter referred to as "Minority" residents or "Minorities") and that this disproportionate effect resulted from racial discrimination by BPD and its officers. The plaintiffs also allege that these discriminatory policies and practices were exacerbated by two significant events: (i) the BPD's creation of a "Strikeforce" unit in June 2012 and (ii) a change in state law shifting responsibility for adjudicating traffic violations issued in the City of Buffalo from the New York State Department of Motor Vehicles to the Buffalo Traffic Violations Agency (BTVA), enabling the City of Buffalo to capture revenue collected from these citations starting in July 2015.

13. The plaintiffs have asked me to provide statistical analyses of the above issues.

III. SUMMARY OF FINDINGS

14. My report analyzes data encompassing the period beginning in January 2012 and ending in December 2022.[1] As discussed below, I separately analyze available data on ticketing ("citations"), stops that did not result in citations ("stops"), and traffic checkpoints conducted by the BPD ("checkpoints).

A. Citations

15. I find that during this period, approximately 75% of citations issued for violations of New York's Vehicle and Traffic Law (not including parking tickets) issued by the BPD were issued to Minority individuals. During the same time period, the population of the City of Buffalo was slightly less than 50% Minority. More precisely, during this time period, Minority individuals in Buffalo received 3.2 times as many citations relative to their population than did Non-Minority individuals relative to their population.

[1] I received TraCS data for 2023 in late March 2024. Given personal and professional commitments, this did not allow me sufficient time to include these data in this Report. I reserve the right to amend or supplement this Report to take account of 2023 data.

16. I find that this racial disparity in citations is much more pronounced with respect to some types of citations than others. In much of my analysis, I separate citations into two categories: (i) Moving violations (which include citations for speeding, reckless driving, failure to stop, failure to signal, driving under the influence, and using mobile devices while driving) and (ii) Non-Moving violations (which include citations for excessively tinted windows, broken lights, worn tires, expired or suspended license or registration, and lacking proof of insurance). I find significant racial disparities (i.e., that Minority individuals were significantly more likely to be cited) with respect to both types of violations. But I find much more pronounced racial disparities for Non-Moving violations. Relative to the sizes of the Minority and Non-Minority populations, Minority individuals received 1.8 times as many Moving violation citations as Non-Minority individuals. By contrast, relative to population size, Minority individuals received over 4 times as many Non-Moving violations as Non-Minority individuals.

17. I find that these racial disparities in citations arise from two sources. First, Minorities received more citations than Non-Minorities relative to their respective populations within neighborhoods. I refer to these disparities in the rates at which Minority and Non-Minority individuals were cited relative to their populations within different neighborhoods as "within-neighborhood" disparities. Second, far more citations were issued in neighborhoods with a higher percentage of Minority residents than were issued in neighborhoods with a lower proportion of Minority residents. I refer to these disparities in the rates at which citations were issued depending on the racial composition of the neighborhoods in which the citations were issued as "across-neighborhood" disparities. As I discuss later, here and throughout the report I define each census tract as a "neighborhood."

18. I find that these across-neighborhood disparities in citations cannot be explained as simply the product of the BPD issuing more tickets in neighborhoods that have greater numbers of criminal incidents or accidents. Instead, even when controlling for neighborhood rates of crime and vehicle accidents, I find that BPD disproportionately issued tickets in predominantly Minority neighborhoods.

Citation Volume Over Time

19. In examining overall ticketing volume over time, I reject the hypothesis that the monthly ticketing process was the same following the implementation of the Strikeforce in June 2012 as it was previously. Instead, monthly citations rates increased significantly after the

implementation of the Strikeforce. I find the ticketing process changed again around the time of the BTVA take-over of ticket revenues in July 2015, with the BPD again significantly increasing the number of citations it was issuing each month. The ticketing process changed again around the time the Strikeforce was disbanded, with monthly ticketing rates decreasing yet remaining well above pre-Strikeforce levels. Finally, I find that monthly ticketing volume changed again around the onset of the COVID-19 pandemic, with monthly citation rates declining to roughly pre-Strikeforce levels and remaining at those levels through 2022.

Racial Disparities in Citations at the Neighborhood Level Over Time

20. These changes in ticketing over time played out very differently in High-Minority neighborhoods than in Low-Minority neighborhoods. Prior to the implementation of the Strikeforce, average monthly citation counts were almost identical in High-Minority census neighborhoods as in Low-Minority census tracts. Following the implementation of the Strikeforce, average monthly citation counts changed little in Low-Minority neighborhoods but more than tripled in High-Minority neighborhoods. Such disparities in average monthly citation counts between High-Minority census neighborhoods and Low-Minority neighborhoods increased even more following the BTVA take-over of the ticketing revenue. After that take-over, High-Minority neighborhoods averaged almost 100 citations issued per month, more than three times the average number of citations issued per month in Low-Minority neighborhoods.

21. These disparities in average monthly citations between High-Minority neighborhoods and Low-Minority neighborhoods lessened following the dissolution of the Strikeforce in early 2018. Even after this dissolution, each High-Minority neighborhood on average still experienced about 45 citations per month until the onset of the COVID-19 pandemic in March of 2020, which was still almost double the number experienced on average by each Low-Minority neighborhood. Following the onset of the COVID-19 pandemic and through the end of 2022, roughly equal numbers of citations were issued per month in High-Minority and Low-Minority neighborhoods.

22. The disparities in the rates at which BPD officers issued citations between High- and Low-Minority neighborhoods in the time period between the implementation of the Strikeforce and the onset of the COVID-19 pandemic cannot be explained as simply reflecting greater issuance of citations in high-crime or high-accident areas. Even when controlling for accidents and criminal incidents, High-Minority neighborhoods still experienced significantly

higher numbers of citations per month than Low-Minority neighborhoods from June 2012 until March 2020. Once again, these citation disparities arose primarily as a result of greater ticketing for Non-Moving violations such as equipment and license/registration infractions rather than as a result of increased ticketing for Moving violations.

Racial Disparities in Citations at the Individual Level Over Time

23. As discussed above, I find that Minority drivers received a disproportionate number of citations relative to Non-Minority drivers over the entire time frame I examine. This disproportionality was present throughout the entire period I studied, but its extent fluctuated over time.

24. Prior to the creation of the Strikeforce in June 2012, Minority drivers received about 1.9 times the number of tickets relative to their population compared to Non-Minority drivers.

25. The racial disproportionality in citations became much greater after the creation of the Strikeforce. After June 2012, Minority drivers received more than three and half times as many citations per capita as Non-Minority drivers. This level of disparity continued until the dissolution of the Strikeforce in early 2018. Even after the dissolution of the Strikeforce, however, Minority drivers continued to receive more than two and a half times as many citations per capita as Non-Minority drivers. This disparity in the rate at which Minority drivers received citations persisted through 2022, the latest year for which I was able to analyze data.

26. As with the neighborhood level disparities, the disparities in the rates at which individual Minority and Non-Minority drivers were cited are greater with respect to Non-Moving violations than Moving violations. Over the period I study, I find that Minorities received between 1.3 and 2.1 times as many Moving citations relative to their population as Non-Minorities. By comparison, during the same period, I find that Minorities received between 2.4 and 4.3 times as many Non-Moving citations relative to their population.

27. As alluded to above, I find that Minorities received disproportionately more citations than Non-Minorities over the entire time frame I examine. I also find that these disproportionalities arose both due to disproportionate ticketing in heavily Minority neighborhoods relative to heavily Non-Minority neighborhoods (i.e., across-neighborhood disparities) and disproportionate ticketing of Minorities relative to Non-Minorities within neighborhoods (i.e., within-neighborhood disparities). However, I find that the contribution of

across-neighborhood disparities relative to within-neighborhood disparities toward the overall racial disproportionalities in ticketing differed over time.

28. Prior to the implementation of the Strikeforce, all of the overall racial disparities in ticketing arose *within* neighborhoods rather than *across* neighborhoods. In other words, prior to the Strikeforce, the BPD issued tickets at fairly similar rates across High- and Low-Minority neighborhoods, meaning the overall racial disparities in ticketing that I find solely reflected disproportional ticketing relative to population within neighborhoods. However, with the implementation of the Strikeforce and the BTVA taking over adjudication, the overall racial disparity in ticketing went up, primarily because of increased disparities in ticketing between High-Minority and Low-Minority neighborhoods during this time frame. In other words, while there remained racially disproportionate ticketing *within* neighborhoods following the implementation of the Strikeforce, in relative terms, these disparities became smaller when compared to the large increase in disproportionalities that arose *across* neighborhoods due to the far heavier citation rates in High-Minority tracts relative to Low-Minority neighborhoods following the implementation of the Strikeforce.

29. Following the dissolution of the Strikeforce and the associated lessening of racial disparities in citation rates between High-Minority and Low-Minority tracts that followed, within-neighborhood rather than across-neighborhood racial disparities in ticketing again accounted for a larger portion of the overall racial disparities in ticketing through 2022.

Racial Disparities in Multiple Ticketing Over Time

30. I find that Minority drivers were significantly more likely to receive multiple tickets within a single incident than Non-Minority drivers during each of the time-frames I analyze. However, these disparities in multiple ticketing are notably greater after the implementation of the Strikeforce. In my analysis of incidents in which drivers were cited for tinted windows, I also find that Minority drivers were significantly more likely to be issued multiple tinted-window tickets than Non-Minority drivers between the creation of the Strikeforce and the start of the COVID-19 pandemic.

B. Stops

31. I also find that between June 2012 and December 2022, Minorities were significantly more likely than Non-Minorities to be issued multiple tickets within a single stop. In

particular, Minorities were statistically more likely than Non-Minorities to receive multiple tinted window tickets in stops in which at least one tinted window ticket was issued.

32. Starting in 2020, the BPD started recording the issuance of stop receipts for stops that did not lead to a citation. I also find racial disparities in this data. Specifically, over three times as many of these stop receipts were issued in the BPD district that had the highest percent Minority population than in the BPD district that had the lowest percent Minority population. Similarly, among the 80% of stop receipts for which driver race was recorded, almost three times as many were issued Minority drivers compared to Non-Minority drivers.

33. Moreover, I also find differences in the reasons given for such stops across BPD districts and across driver race. Notably, in BPD districts with higher Minority populations, a significantly larger fraction of stops that did not lead to tickets were for equipment issues or for failure to use a turn signal. By contrast, in districts with smaller Minority populations, a larger fraction of such stops were for failure to stop at a stop sign or stop light or erratic/reckless driving. Results were similar when looking at reason for stop by race of driver.

C. Checkpoints

34. I also find the BPD overwhelmingly located checkpoints in heavily Minority neighborhoods relative to lower Minority neighborhoods. Indeed, 60% of neighborhoods that were less than 40% Minority (which I refer to as “Low-Minority”) experienced fewer than 5 checkpoints total, and none experienced more than 30. By contrast, fewer than 15% of neighborhoods that were more than 60% Minority (“High-Minority”) experienced fewer than five checkpoints, while over 35% experienced 30 or more, with almost 17% experiencing more than 75 checkpoints. Such discrepancies in checkpoints between High-Minority and Low-Minority neighborhoods cannot be explained by the BPD simply locating checkpoints in the neighborhoods with larger populations or with higher levels of crime or accidents.

IV. DATA SOURCES AND MATERIALS CONSIDERED

35. This section of the report describes the data sources on which I relied in conducting my analysis.

36. In preparing this report, I considered a variety of materials, including the pleadings, discovery responses, documents produced by the parties and third parties, deposition transcripts, and various other materials. Exhibit B hereto is a list of materials considered. I also relied on my knowledge and experience in the fields of statistics, racial discrimination, and policing.

A. Citations Data

1. TraCS

37. TraCS is the primary dataset used for my analysis and contains information on all motor vehicle citations given in the City of Buffalo between January 1, 2012, and December 31, 2022.[2] This dataset was obtained via subpoena from Erie County Central Police Services (ECCPS), which maintains it on behalf of the City of Buffalo. Counsel obtained field descriptions, and other information concerning the dataset was obtained for me from ECCPS personnel.

38. I received the TraCS data in four different tranches at different times. The original tranche contained data through March 2019; subsequent updates brought this data up to date through December 2022.[3] The initial tranche contained anonymized identifiers in lieu of actual Driver's License (DL) numbers, but the later tranches contained actual DL numbers, and I received a concordance so that I was able to use actual DL numbers in all cases where that information was originally present.

39. These tranches of data were merged into one dataset. Of all the observations in this dataset, I kept only the ones that were issued by the Buffalo Police Department. There was some overlap across tranches that led to some duplicate observations. I identified these duplicate observations based on citation number, date, and time (to the minute); where there were duplicate observations, I kept only one. There were also several citations that had the same citation number but appear to be distinct citations in that their date, time, location, and violation codes were distinct. I kept all of these observations as they appear to be distinct citations. In total, the TraCS dataset contains information on 374,386 distinct citations.

40. For all citations in the TraCS dataset, there was information regarding the date of issuance, and for all but one citation, there was also time (to the minute) of the issuance. This data set also contained violation codes associated with the citation for all but 1,205 observations. Using these violation codes along with the New York Vehicle and Traffic Laws, I categorized citations into two subcategories: (i) Moving violations (including speeding, reckless driving, failure to stop,

---

[2] The dataset as supplied by ECCPS actually goes back to 2011, which is when the BPD started transitioning from paper to electronic ticketing. However, the 2011 data appear to be incomplete. Accordingly, I treat the starting point of the dataset as January 1, 2012.

[3] As noted above, I recently received TraCS data for 2023, but these data were produced too late to be included in this Report. *See supra* note 1.

failure to dim headlights, DUI, improper use of mobile communications device), or (ii) Non-Moving violations (including equipment violations such as tinted windows, broken windows or lights, obscured license plate, lack of or improper use of seatbelts or car seats, obstructed view, bald tires, or loud muffler, as well as license and registration violations, insurance violations, and inspection violations). There were just over 5,000 citations that did not fit in either of these categories, as they consisted of Transportation Law infractions by commercial vehicles.

41. Some of my analyses consider the neighborhoods in which citations were issued. For these analyses, I use a census tract to define a neighborhood. A census tract is an area determined by the U.S. Bureau of Census that is roughly equivalent to a neighborhood, with the area in each tract being contiguous and the boundaries generally following visible and identifiable features.[4] Census tracts have an optimum size of 4,000, but can range from a little over 1,000 to close to 8,000 inhabitants. During the period I studied, census tracts in the City of Buffalo generally encompassed about 3,500 residents.

42. TraCS data was provided by Erie County in a format that made it somewhat difficult to convert location information into geolocated data. The reason is that location information, which is contained in the "vio_street" variable, is a written description provided by the citing officer of where a citation was being issued. Thus, extensive effort was required to convert written descriptions, which often included shorthand descriptions, misspellings, abbreviations, and the like, into usable location descriptions. This process was completed through several steps, which are outlined below. This process of organizing and coding data is generally referred to as "cleaning" the data and is a necessary and standard step in almost any type of data analysis.

43. Much of this cleaning and coding was performed by Dr. Gregory DeAngelo, an Associate Professor of economic sciences and the Director of the Computational Justice Lab at the Claremont Graduate University. Dr. DeAngelo has extensive experience in this type of work, and is, in my view, highly qualified to perform this work. In cleaning and coding this data, Dr. DeAngelo worked under my direction and supervision. I reviewed and tested the results of his cleaning and coding of data and periodically met with him via videoconference to discuss his work.

---

[4] Census Bureau, Glossary: Census Tract, https://www.census.gov/programs-surveys/geography/about/glossary.html.

In my experience, it is common to use data that has been cleaned by collaborating researchers or research assistants working under one's direction and supervision.

44. The first cleaning step was to remove observations that were not associated with "BPD" or "Buffalo" entries in the "police_agency" variable. This step ensured that all data analyzed pertained exclusively to citations by the Buffalo Police Department.

45. Second, a series of steps were taken to identify information about the location of the citation based on the information contained in the "vio_street" variable. To start, the "vio_street" variable was cleaned to determine if the location of the citation involved a numerical digit. This often, but not always, indicated that a citation was issued at a specific street address. In this case, attempts were made to extract the street number and name so that the geolocation could be identified based on a street address. Next, numerous extractions of the directionality (e.g., north, south, northeast, etc.) of the street where the citation was issued were conducted. Code was then written to identify when words associated with a street were existed in the "vio_street" variable. These words included "street", "st.", "avenue", "ave", "rd", etc., which were extracted from the "vio_street" variable. Finally, the actual street name was extracted and assumed to be the remaining information in the "vio_street" variable that is not associated with a numerical value, street label, or directionality.

46. Based on these cleaning steps, it was determined that approximately 35% of the observations in the TraCS data involved a street address, while approximately 63% of the observations involved an intersection of two street names.

47. Once the raw TraCS data were cleaned so that each observation was determined to either be a street address or an intersection (when possible), they were geocoded using Google's geolocation application programming interface ("API"). To convert addresses into geographical coordinates, a geocoding API developed by Google was utilized. Geocoding is the process of transforming a description of a location—here, a street address or intersection—to a location on the earth's surface. To improve the geocoding process, information related to the location, such as the address, city, county, and state, are added to the location information so that the Google API can more easily understand the address. This step improves the accuracy of the geocoding process. The location information is then loaded into a Google developer console that enables geocoding of the location. This process generates an API key required to obtain the

geographic coordinates associated with each address. The complete address information was then passed through the API key, which produced a latitude and longitude associated with each address.

48. Within the TraCS data, we were able to geocode approximately 98% of the 374,386 observations. With respect to the 2% of observations we were unable to geocode, approximately 15% had no address information contained in the "vio_street" variable, with the remainder having address information that was not specific enough to identify a location (e.g., "Delaware Ave", "Doat St", "West").

49. The final step to fully geocode the TraCS data was to attach census tract information to each observation, as census tracts will constitute my definition of a neighborhood and often used as my unit of analysis. Since the previous step attached latitude and longitude information to more than 98% of the data, this final step simply required identifying which census tract corresponded to each set of geographical coordinates.[5] For this step, the obtained coordinates were then passed into an open-source Census conversion API[6] to obtain the census tract associated with each location that is based on the latitude and longitude information obtained from the Google API. This process was 100% successful for observations where latitude and longitude information were available. While this process is successful at associating locations with census tracts, there are rare instances in which a single location can be mapped into two census tracts. This occurs in less than 0.5% of all observations and typically occurs when a location is near the boundary of two census tracts. For the small number of instances where the same or similar addresses were assigned to different census tracts (likely due to adjustments in Google's geocoder or the Census API), my team assigned each address to the census tract to which that address was most often mapped.

50. Using this procedure, we were able to map 366,618 observations to census tracts, with over 99.7% of these observations mapped to census tracts within the City of Buffalo. Given that I analyze citation, stop, and checkpoint activity at the neighborhood level, and define each census tract as its own neighborhood, for all neighborhood-level analyses, I use only those observations that could be mapped to City of Buffalo census tracts.

---

[5] For consistency, I use the 2010 census tract boundaries throughout my analysis. These were unchanged until the 2020 Census. At that time, an handful of census tracts in Buffalo were split or otherwise altered.

[6] More information about the Census's geocoding API can be found here: geocoding.geo.census.gov/geocoder/Geocoding_Services_API.html.

51. The TraCS dataset contained a non-blank value for Driver's License (DL) number for over 98% of the citations. However, for just over 2000 of these citations, the value under the driver's license variable does not appear to be a valid DL number (e.g., it is listed as "UNKNOWN" or "000000" or "N/A"), meaning we have DL numbers for 97.6% of the records in TraCS.

52. In addition to date, time, location, law section, and DL number, the vast majority of TraCS observations also contains the zip code of the driver's age, sex, and most recent residence.

53. Using some of the above variables, I was also able to identify distinct incidents in the TraCS data via the following process. I first defined a TraCS incident to include all citations issued on the same date, same time (to the minute) to someone with the same DL number. For observations with associated DL numbers, I defined the same incident to include all citations with the same location, same date, same time (to the minute), same sex, and same age. Among those observations missing license number and sex, I then defined a TraCS incident to include all citations with the same location, same date, same time (to the minute), and same age. Among the remaining observations missing license number and age, I then defined a TraCS incident to include all citations with the same location, same date, same time (to the minute), and same sex.

54. In the end, the 374,386 citations in TraCS were allotted to just over 160,000 separate incidents, meaning many incidents resulted in more than one citation.

55. Finally, the TraCS data contains information on race of cited individual for some observations. As I discuss in more detail in the Methodology section below, I merge in data from the Open Data Buffalo dataset to check and augment this race information contained in TraCS.

2. Open Data (OD) Buffalo

56. The City of Buffalo posts data regarding a variety of police activities including individual citations issued in the city via the website https://data.buffalony.gov/. These data are downloadable. Through counsel, I have been informed that the source of the data is the SUMMONS table from the CHARMS system, described below. Much of the SUMMONS data is derived directly from TraCS, but the SUMMONS table also has access to the CHARMS MASTER_NAME table and thus to demographic information (in particular, race and ethnicity) separate and apart from that recorded in TraCS. As discussed below, when I am able to link observations in the TraCS dataset to the OD data set, race and ethnicity across these two sources

almost always agree when both are present. However, there are many TraCS observations in which race/ethnicity is not recorded, but can be obtained from matched incidents from OD, allowing me to increase the overall number of observations for which race/ethnicity of cited driver is documented.

57. Each observation in this OD dataset is associated with a citation and contains information regarding date and time of the citation. For many citations, the observation also includes information about the age, sex, and race of the cited individual. I used this dataset primarily to (i) check the race information contained in the TraCS data, and (ii) fill in race information for citations in which TraCS did not contain race information for the cited individual.

58. In order to match OD data to observations in TraCS, I defined OD incidents as follows. I first defined an OD incident to include all citations with the same location, same date, same time (to the minute), same sex, and same age. Among those OD observations missing sex, I defined an OD incident to include all citations with the same location, same date, same time (to the minute), and same age. Among the OD observations missing age, I defined an OD incident to include all citations with the same location, same date, same time (to the minute), and same sex. Among the 34 OD observations missing both age and sex, I defined an OD incident to include all citations with the same location, same date, same time (to the minute). Because OD data does not contain any sort of individual identifier such as DL number, I could not start with incidents defined as all citations occurring at same date and time with the same DL number (or other personal identifier) as I did in TraCS.

59. I then made two data sets, one containing TraCS incidents (including date, time, gender, and age) and another containing OD incidents (including date, time, sex, and age).[7] I then matched TraCS incidents to OD incidents based on having the identical date, time, age, and gender (when gender was not missing) or date, time, and age (when sex was missing). Given this procedure, it was not possible to match TraCS incidents to OD incidents for incidents that were not unique in terms of date, time, age, and gender (or date, time, and age when gender was

[7] There were 12 observations in the OD dataset in which the race measure was inconsistent within the assigned incident, meaning some of the citations in the incident were categorized as being issued to a Black driver, while others were categorized as going to a non-Black driver. . For these incidents, I changed the race measure to missing. Similarly, there were 4 observations in the OD dataset in which the Hispanic ethnicity measure was inconsistent within the assigned incident. For these incidents, I changed the Hispanic ethnicity measure to missing.

missing). Therefore, any such incidents could not be matched between the TraCS and OD data, meaning we could not use the OD data to validate or fill in missing race data for any such TraCS incidents.

60. In the end, I was able to match over 95% of the TraCS observations that were mapped to a Buffalo census tract. Among the TraCS incidents that were matched to a unique OD incident, all observations within that incident were matched to the OD information on driver race associated with that incident.

61. I discuss further details of the race measure obtained from OD Data, as well as the race measure contained in the TraCS data, in my analysis of cited drivers below.

B. Checkpoint Data

62. Data on checkpoints was received in several ways. From a Freedom of Information Law request, counsel obtained and subsequently passed on to me a listing of what the BPD asserted were all of the checkpoints operated by the Strikeforce, together with identification of the census tract(s) of the checkpoints. When checkpoints were on the border of two or more census tracts, all tracts were identified. In cases in which a checkpoint was associated with multiple census tracts, I "divided" such checkpoints across the listed census tracts. So, for example, if a checkpoint had two census tracts associated with it, I allocated one half of a checkpoint to each of these census tracts. In cases in which a checkpoint had three census tracts associated with it, I allocated one-third of a checkpoint to each of those census tracts. This approach avoids double- or triple-counting checkpoints and means that my count of checkpoints will be consistent even when I aggregate to the census tract level.

63. During the course of discovery, this listing was updated to include the actual street address or (more frequently) the intersection at which the checkpoint was located. The defendants also produced documents known as "Roadblock Directives," which contained information on location, date, and time of each checkpoint. Counsel were informed that the previously produced listing of checkpoints was derived from these roadblock directives.

64. Finally, it is not clear that every checkpoint on the list produced by the defendants and derived from the Roadblock Directives necessarily took place. I understand that there are approximately 373 listed checkpoints in which no tickets were issued in the listed Census tracts within an hour of the listed starting time of the checkpoint. Therefore, I also assess the robustness of my results if these checkpoints are excluded from the analysis.

C. Stops Data

65. Beginning in late June 2020, the BPD began issuing "Traffic Stop Receipts" (TSRs) when an individual was pulled over but it was ultimately decided not to issue a citation. BPD directives indicated that every traffic stop should result in either a citation or a TSR (*i.e.*, no stops should remain undocumented).[8] My understanding is that those directives were not always adhered to. In particular, counsel has informed me that there are numerous events characterized as TRAFFIC STOP in the CHARMS database that do not connect to either a TSR or a citation. For this reason, my use of the term "Traffic Stop" (or simply "Stop") is limited to stops in which a TSR was issued.

66. My understanding is that officers issued TSRs using the ENTMobile device in their cars, which was a separate mobile electronic device from the TraCS interface.[9] I understand that ENTMobile fed the CHARMS system. Counsel obtained TSR data from ECCPS via subpoena.

67. The TSR data contains race and/or ethnicity of driver for some observations, but this data is incomplete. As I discuss in more detail below, for TraCS observations that were missing race/ethnicity information and could not be matched to an OD observation that contained such information, I use the zip code of residence along with other observable information to impute the racial distribution for citations with unrecorded race for some analyses. However, the original TSR data did not contain address or zip code for the drivers receiving TSRs. A supplemental production joined the TSR table with the CHARMS MASTER_NAME table, which then provided addresses (including zip codes) for some but not all of the drivers.[10] This zip code of residence information remains sufficiently incomplete that I cannot use it to impute information about driver race, as I do for the TraCS data. Therefore, in my analysis of the TSRs, I treat TSRs with unrecorded race and/or ethnicity as a separate category.

D. Neighborhood Demographic Data

68. I use data from the American Community Survey (ACS) to obtain demographic information about the residents of the City of Buffalo by neighborhood.

---

[8] Training Bulletin, Ent Mobile – Electronic Traffic Stop Receipt, COB560603.
[9] *Id.*
[10] Specifically, drivers who had been the subject of a previous COMPLAINT recorded in CHARMS would have entries in the MASTER_NAME table, and those entries frequently included Race and/or Ethnicity.

69. The ACS is a nationwide survey of households each year conducted by the United States Census Bureau. It collects information on economic, demographic, and housing characteristics of the United States population every year. For each of the years between 2012 and 2019, the ACS collected information from approximately 2 million housing units nationwide. The sample is structured so that the ACS data are representative of the United States population.[11]

70. The primary use of the ACS data in my analysis is to determine the racial composition of each neighborhood (i.e., census tract) in the City of Buffalo. The ACS is generally considered the best available source for capturing racial demographic information for states, counties, cities, and census tracts for any given year. However, even though the ACS is an annual survey, I do not think it is appropriate to use it to capture changes in the racial composition of units as small as census tracts over time for the following reasons. As stated above, the ACS is a representative sample of the U.S. population each year, not a census. Therefore, in the ACS for any given year, the measured number of residents of any given race (including the total number of residents) in each census tract is an estimate based off a five-year survey average. Accordingly, the number of residents in each census tract across years fluctuates due both to residents moving in and out, as well as to measurement error because of sampling variation and averaging over several years. An additional complication to using the ACS to measure the racial composition of each census tract in each year is that a new decennial census was conducted in 2020. As part of each decennial census, the Census Bureau changes some census tract boundaries, meaning the boundaries of census tracts in the City of Buffalo in 2012-2019 may not correspond to census tract boundaries from 2020 onwards.

71. Given these issues, I measure the racial composition of each census tract as follows. From the ACS, I obtain the annual estimates of the total population, the total Black non-Hispanic population, and the total Hispanic population for each census tract in Erie County for the years 2012-2019 (as noted above, these measures are actually themselves based off of surveys over the previous five years). I then select only the 79 census tracts that are contained inside the City of Buffalo (using 2010 census tract boundary definitions). Using these data, I calculate the fraction of each Buffalo census tract that is Black and/or Hispanic for each ACS year. Note that I combine

[11] For background about the ACS, see Census Bureau, American Community Survey: Information Guide (Oct. 2017), and Census Bureau, About the ACS, https://www.census.gov/programs-surveys/acs/about.html (June 27, 2023).

Black and Hispanic residents together as "Minority" residents, as the Black and Hispanic categories are not mutually exclusive.[12] Therefore, to avoid double counting, I measure the fraction Minority of each census tract as the number of people who identify as Black/Non-Hispanic plus the number of people who identify as Hispanic, divided by the total number of people in the tract. I then average this measure over the 2012 to 2019 timeframe for each census tract to come up with one value for each census tract capturing the average percentage of residents of that census tract who were either Black or Hispanic (i.e., Minority) between the years 2012-2019.

72. This approach gives me the most accurate possible measure of the fraction Minority on-average for each census tract over the 2012-2022 period examined here. Although there is some measurement error for any given year, such measurement error likely has very little effect on my analysis because in most of my neighborhood-level analyses, I use a coarser way to categorize the racial make-up of census tracts, and this coarser categorization should be minimally impacted by this measurement error. Namely, given my measure of the fraction Minority for each Buffalo census tract as described above, I generally categorize each census tract as High-Minority (defined measured to be 60% or more Black and/or Hispanic on-average between 2012 and 2019), Low-Minority (defined to be less than 40% Black and/or Hispanic on-average between 2012 and 2019), or Mixed-Race (defined to be 40-60% Black and/or Hispanic on-average between 2012 and 2019).

73. The reason this classification system mitigates measurement error is that even though the fraction Minority of each census tract certainly changes over the time period examined in this analysis, it is unlikely that these changes would often (if ever) be so large that tracts would actually change categories within the span of the time period examined here. Tracts are especially unlikely to flip from being High-Minority to being Low-Minority (or vice versa). Indeed, while the annual ACS measures are estimates, and therefore subject to measurement error as discussed above, there exists only one tract that would be classified as a High-Minority tract in some years and a Low-Minority tract in other years on the basis of annual ACS estimates if tracts were separately classified each year. On average over the years 2012-2019, this one tract had a fraction Minority right around 50%, and therefore is defined as a Mixed-Race tract in my analyses. This

[12] The ACS has separate questions regarding race and ethnicity, since an individual can be Hispanic and Black or Hispanic but not Black. According to the ACS, about 8% of Buffalo residents who classify themselves as Black also classify themselves as Hispanic.

means this tract will not have any impact on my key analyses in which I focus primarily on comparing High-Minority tracts to Low-Minority tracts.

E. Additional Data

1. Data on Motor Vehicle Accidents

74. Information on accidents was initially obtained through mid-2019 from a Freedom of Information Law Request to the New York State Department of Transportation. Subsequently, my team was able to obtain 2018–2022 data using the Department of Transportation's website. Key information available for each accident includes the date and the accident, whether there was an injury, whether alcohol or drugs were involved, and location.

75. Using this location information, we were able to map each accident to a census tract using similar methods as described earlier.

2. Data on Crime

76. Data on crime comes from the Complaint History and Record Management System (CHARMS) maintained by ECCPS on behalf of the City of Buffalo and was obtained by counsel via subpoena to ECCPS. As in the case of TraCS, several updates of the CHARMS information were received over time, and the final dataset used in this report covers the period from 2012 through December 2022.[13] ECCPS provided information concerning the structure of CHARMS to counsel by ECCPS, who in turn relayed that information to me as relevant and when requested to do so.

77. CHARMS consists of a series of linked data tables. The entry-level table is the COMPLAINT table; the entries in the INCIDENT, ARREST, and OFFENSE tables all have a COMPLAINT number and can be traced back to the COMPLAINT table. There are also ancillary tables, such as the MASTER_NAME table, which contains such biographic and demographic information as was gathered concerning individuals who have been the subject of a COMPLAINT.

78. For the purposes of my analysis, "crimes" are defined by what the BPD categorizes as criminal incidents (i.e., entries in the INCIDENT table). Within this category, I have categorized incidents into violent crimes (e.g., murder, kidnapping, robbery, assault, rape and other sex offenses, dangerous weapons, offenses against family) and property crimes (arson, burglary,

[13] As noted above, additional data through the end of 2023 were received too late to be incorporated in this report.

extortion, forgery and counterfeiting, fraud, larceny, motor vehicle theft, possession of burglar tools, stolen property, unauthorized use of vehicle).

79. Each criminal incident had location information sufficient to map it to a census tract in the manner described above using the TraCS data of tickets. Given this mapping, along with dates of the incidents and the type of crime, I was able to create counts of violent crime incidents and property crime incidents arising in each census tract in different time periods.

V. EMPIRICAL METHODS

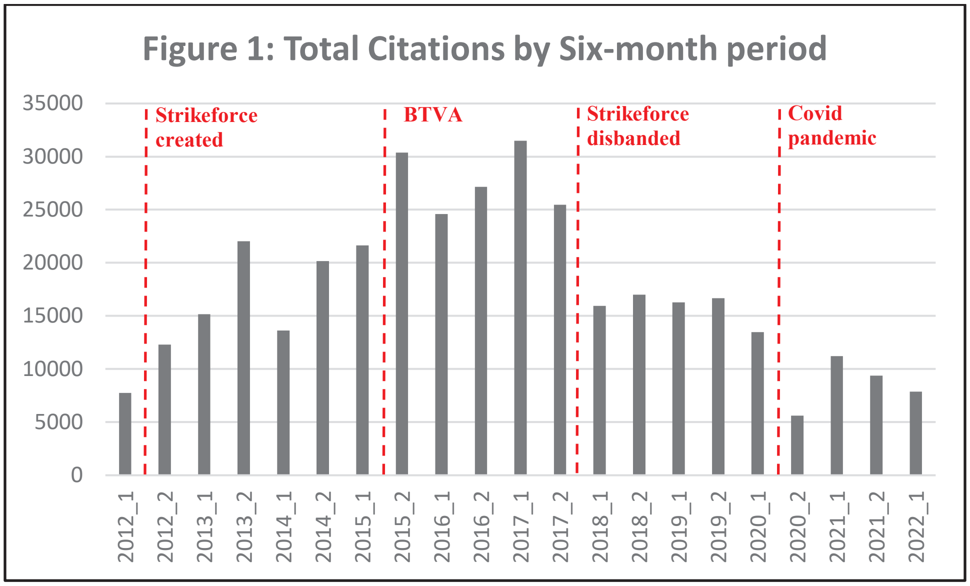
80. The sub-sections below detail the principal empirical approaches for analyzing the data described above. Each of the methods I describe below is a common and well-accepted method of statistical analysis.

A. Methods for Analyzing Citation Activity Over Time

81. The plaintiffs allege that the BPD increased its issuance of citations following the implementation of the Strikeforce in June 2012, as well as following the BTVA take-over of ticketing (and associated revenue) for the City of Buffalo in July 2015.

82. To assess this claim, I focus on evaluating the extent to which there are measurable and significant changes in citation counts corresponding to the creation of the Strikeforce and the BTVA take-over of ticketing mentioned above. In addition to those two events, I also consider other events that may also have had notable impacts on BPD citation practices. First, in December 2017, following a resolution issued by the Buffalo Common Council accusing the BPD of racial discrimination and the filing of a civil rights complaint by Black Lives Matter-Buffalo, the New York State Attorney General opened an investigation into racially biased citation practices by the BPD. This was followed closely by the retirement of BPD Commissioner Derenda in January 2018 and by the Strikeforce being disbanded in early February 2018. Second, the COVID-19 pandemic began in the United States in March 2020, shutting down businesses and schools. In May of 2020, George Floyd was killed by Minneapolis Police officers, setting off protests throughout the United States, including in Buffalo.

83. Figure 1 shows the overall number of citations issued by the BPD between 2012 and the end of 2022 by six-month period. The vertical lines are approximately associated with the events described above.[14]



84. As can be seen in Figure 1, it is clear that there are large changes over time in the number of citations issued in each six-month period. At least in broad terms, some of these changes appear to coincide with the dates of the key events described above.

85. For much of my analysis, I will divide time into five distinct time eras: (i) the Pre-Strikeforce Era (January 2012 – June 2012), (ii) the Strikeforce Era (July 2012 – June 2015), (iii) the BTVA Era (July 2015 – January 2018), (iv) the Post-Strikeforce Era (February 2018 – February 2020), (v) the Current Era (March 2020 onward). I use these era names solely as a short-hand for ease of reference, with the chosen date ranges meant to approximately line up with several key events. While these time eras roughly line up with the key dates discussed above, it is not necessarily the case that changes in BPD behavior line up precisely with these dates. For example, there may have been some lag time between when the Strikeforce was announced and when it began to conduct policing activities, and the Strikeforce may have started ramping down its activities before it was officially disbanded. However, my aim is to analyze whether general

[14] As noted, the Strikeforce was not officially disbanded until early February 2018 (one month into the first bar of 2018), and the Covid-19 pandemic did not trigger lockdowns and other changes in behavior that might be expected to influence citation activity until March 2020 (three months in to the first bar of 2020).

citation practices differed across the time eras defined above, where the eras may generally reflect different states of the world.

86. In evaluating whether citation practices changed in a statistically significant manner across these eras, my variable of analysis is monthly citation counts. Because the Pre-Strikeforce Era is only six months long, I do not want to use statistical methods that rely on large sample sizes. Moreover, I want to impose as few assumptions as possible on underlying nature of the data generation process within or across eras.

87. I therefore use a Wilcoxon Rank-Sum test, which is also sometimes referred to as a Mann-Whitney test. This provides a way to test whether two samples come from the same underlying data generation process. It does so without making any assumptions on the distributional form of the data generation process, nor does it require particularly large sample sizes for accurate hypothesis testing.

88. The basic idea of this test is as follows. For illustrative purposes, suppose we have sample *a* consisting of N observations of numeric values drawn from some data generation process *A* and sample *b* also consisting of N observations of some numeric values drawn from some data generation process *B*. The goal is to use these samples to assess whether the data generation process *B* is different than *A*. To make such an assessment, the Wilcoxon Rank-Sum test first combines the two samples together (but keeps track of which observation is from which sample) and then ranks these observations from the lowest value to the highest value. The test then separately sums the rank of all the observations from sample *a* and then does the same for the observations from sample *b*. If the two samples come from the same underlying data generation process, then we should expect this sum of ranks should be roughly equal between those coming from sample *a* and those coming from sample *b*. Alternatively, if the sum of ranks is much lower than expected for those coming from one sample than the other, this would constitute evidence against the hypothesis that the two samples come from the same underlying data generation process, and therefore in favor of the alternative hypothesis that the two samples come from different underlying data generation processes. This is the basic idea behind the Wilcoxon Rank-Sum test. The actual test, however, does not need the samples to be of the same size N.[15]

[15] Formal descriptions of the Wilcoxon Rank-Sum test can be found in most upper-level statistics textbooks, for example Devore, Jay. 1995. Probability and Statistics for Engineering and the

89. In the context of this evaluation, each era j is treated as a "sample" of N_j observations, where N_j is the number of months in era j. Each observation in a sample is the number of citations in a month, for each of the N_j months in that era. I then use a Wilcoxon Rank-Sum test to assess the null hypothesis that the monthly citation counts in any two adjoining eras come from the same underlying data generation process. Statistically rejecting this null hypothesis constitutes evidence that the two eras have different data generating processes with respect to monthly citations. In other words, absent large changes in civilian behavior between any two adjoining eras, a rejection of the null hypothesis would constitute evidence that policing practices with respect to citations differ across the two eras.

90. Given that the variance in citation counts within any era is not trivial (as can be seen in Figure 1), one might worry that this Wilcoxon Rank-Sum test errs on the side of rejecting the null hypothesis that monthly citation counts across any two adjoining eras come from the same underlying process even if they actually do. To assess whether this concern is valid, I also perform a type of "placebo" exercise in which I randomly divide the months within each era into two subsamples, then use the Wilcoxon Rank-Sum test to assess whether the monthly citation counts in these random subsamples of months from within a given era come from the same underlying data generation process. I do this randomization 10 times for each era. If the Wilcoxon Rank-sum test is not overly sensitive (i.e., that it frequently falsely rejects true null hypotheses), few of these within-era tests should reject the null hypothesis that those within-era samples come from the same underlying data generation process.

B. Methods for Analyzing Racial Disparities In Citations

91. The second question I analyze relates to whether the BPD's citation practices have disproportionately affected Black and Hispanic residents of Buffalo. In analyzing this question, I look at how citation practices may differ by the racial demographics of the neighborhood of issuance, as well as by the race of individual drivers. I analyze racial disparities with respect to citation practices separately for the five eras described above.

92. In my analysis, I employ a variety of different approaches in order to avoid relying on a particular inclusion criteria or specific assumptions. Employing these different methodological approaches allows me to assess whether these different approaches yield

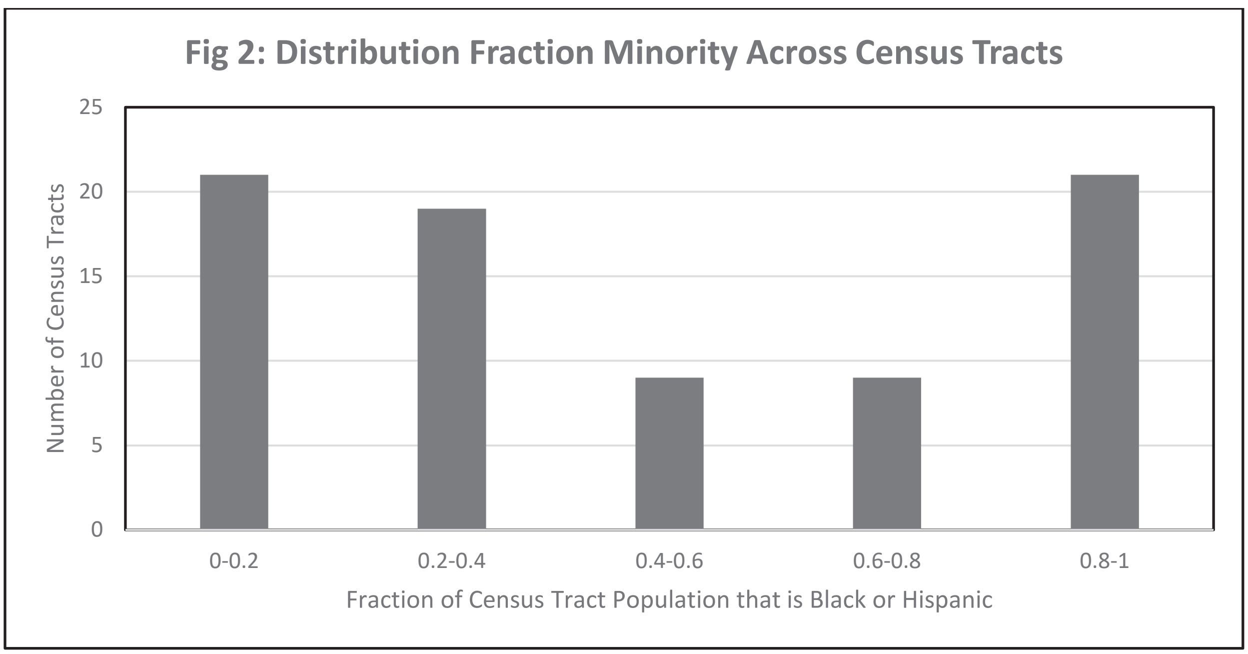
Sciences, Fourth Edition. Duxbury Press. For my estimations, I use the "ranksum" package in Stata 17.

consistent findings regarding whether BPD citation practices disproportionately affected Minority residents of Buffalo.

1. Methods for Analyzing Racial Disparities in Citations Across Neighborhoods

93. As discussed earlier, I use census tracts as my definition of neighborhoods for analysis purposes. Census tracts are defined by the Census Bureau. These are meant to have a population size of between 1,200 and 8,000 people, with an optimum size of 4,000 people, and not be broken up by major obstacles to movement such as thoroughfares, highways, or rivers.[16] Given their precise and relatively consistent boundaries, census tracts are frequently used as the measure of neighborhoods for statistical analysis purposes.

94. Figure 2 shows the distribution of census tracts in the City of Buffalo by the percentage of the residents of that tract that are classed as Minority, using the ACS data described above. This figure reveals that Buffalo exhibited substantial racial segregation over the time period examined in this analysis, with most of its citizens living in quite racially homogeneous neighborhoods as captured by Census tracts. Indeed, Figure 2 shows that almost half of Buffalo residents live in census tracts that are over 80% Minority or 80% Non-Minority.[17]



[16] https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_13

[17] According to the ACS, about 8% of Buffalo residents who classify themselves as Black also classify themselves as Hispanic.

95. In much of my neighborhood-level analysis I characterize each census tract as being: (i) High-Minority (60% or more of the population is Black and/or Hispanic), (ii) Mixed-Race (between 40% and 60% of the population is Black and/or Hispanic), or (iii) Low-Minority (40% or less of the population is Black and/or Hispanic). Looking again at Figure 2, given these definitions, we can see that 30 of Buffalo's 79 census tracts are classified as High-Minority (the two rightmost columns in Figure 2), 40 are classified as Low-Minority (the two leftmost columns in Figure 2), and just 9 are classified as Mixed-Race (the middle column).

96. Among the tracts classified as High-Minority, the average fraction Black and/or Hispanic is over 83%. By contrast, among those tracts classified as Low-Minority, the average fraction Black and/or Hispanic is less than 22%. The average fraction Black and/or Hispanic in those census tracts classified as Mixed-Race is approximately 50%. On average, tract population is almost identical across these three tract classifications.

97. I focus most of my analysis on comparing BPD citation practices in census tracts categorized as High-Minority relative to census tracts categorized as Low-Minority for several reasons. First, as discussed above, tracts that are classified as High-Minority are very distinct with respect to their racial demographics when compared to tracts classified as Low-Minority. Second, while it is true that the racial composition of each census tract changes over time, as discussed in Section III.C.1, the ACS data is collected from only a sub-sample of the American population each year. This means that variation in things like racial composition of each census tract across years is a combination of true changes and sampling variation. As discussed earlier, however, any tracts whose variation in racial composition is such that it would be classified differently in different years would generally be classified as a Mixed-Race tract using my methods here.

98. My focus is on comparing citation counts in High-Minority tracts relative to Low-Minority tracts, given these are the two categories that make up the vast majority of tracts in Buffalo and most clearly differ in terms of their racial demographics. For completeness and transparency, however, I also include results for those tracts classified as Mixed-Race when my methods allow.

99. Determining what cutoffs to use to define these three categories of census tracts, and even the choice to divide tracts into three categories rather than two or more than three, is not an exact science. In making this choice, my goal was to balance a few objectives. First, I want the tracts within each category to be very distinct from the tracts in the other categories with respect

to racial demographics. Second, within each category, I also wanted to have enough tracts to allow for statistical inference. This guided me to focus on the High-Minority (n=30) and Low-Minority (n=40) categories, but also include a third Mixed-Race category (n=9) to ensure that the High-Minority and Low-Minority categories were sufficiently distinct when it came to racial demographics. Mean tract population is quite similar across all three of these tract categories. Comparing just the High-Minority and Low-Minority tracts, I also find that the 10th, 25th, 50th, 75th, and 90th percentiles of the population distribution are all within a couple of hundred people across these two categories (with Low-Minority tracts generally being slightly more populous).

100. While conducting this analysis, I experimented with slightly different cutoff definitions for these categories, and the precise choice of cutoffs did not meaningfully affect my results. Indeed, in analyzing whether there are racial disparities in BPD citation practices across neighborhoods, my starting point is to simply assess whether there are large disparities in citations by tract racial demographics. For this more descriptive analysis, I divide up census tracts based on racial composition into additional more narrowly defined categories than just the three above. In this way, I am able to visually represent what types of heterogeneities there are even within the High-Minority, Mixed-Race, and Low-Minority categories discussed earlier.

101. It is important to note that, to the extent there are disparities in average citation counts across these tract type categories, these disparities may not necessarily reflect racial bias in policing for several broad reasons. First, it may be that the prevalence of violations simply differs across these tract categories. If this is true, then even if police employ identical citation practices in High-Minority tracts as in Low-Minority tracts, there would be more citations issued in High-Minority tracts than Low-Minority tracts. I cannot evaluate this argument directly, as I do not have data on the underlying prevalence of violations by census tract nor by race of driver (nor do I think such data is attainable for many if not most types of violations). However, to at least partially address this issue in the neighborhood of issuance analysis, I analyze citation disparities between High-Minority and Low-Minority tracts separately across the five different eras described earlier. I do so because it would seem very unlikely that the underlying prevalence of violations in any neighborhood would change abruptly around the key events that delineate the five eras described above. Therefore, any abrupt changes in citation disparities between High-Minority and Low-

Minority tracts around these events would suggest a change in citation practices by the BPD that affected High-Minority tracts differently than Low-Minority tracts.[18]

102. The second way in which any citation disparities between High-Minority and Low-Minority tracts may not reflect a racial bias in policing is if such disparities arose because policing practices focused on things other than the racial make-up of the neighborhood, but which happened to be correlated with the racial composition of the tract. For example, if the BPD chose to deploy more officers to areas with higher crime, police might have issued more citations in these locations, as there would have been more police to observe violations occurring in those areas. If these higher-crime areas were more prevalent in High-Minority tracts than in Low-Minority tracts, this refocus of policing activity could have indirectly led to more citation activity in High-Minority tracts relative to Low-Minority tracts. Similarly, there could be disparities in accidents between High-Minority and Low-Minority tracts.

103. I account for the above possibility using three different methodological approaches: (a) Analysis of High Violent Crime tracts only, (b) Multi-Variate Linear Prediction, and (c) Propensity Score Matching. I discuss each of these approaches below.

a) Analysis of High Violent Crime Tracts Only

104. The first approach is to do a sub-analysis of just *High Violent Crime* (HVC) tracts, where I define such tracts to be those tracts in the highest third of the distribution with respect to *violent crime incidents* in the 12 months leading up to the implementation of the Strikeforce. I then compare average monthly citation counts in these HVC High-Minority tracts relative to the HVC Low-Minority tracts in each of the different eras and use a t-test to assess whether monthly citation counts differ between the High-Minority HVC tracts and the Low-Minority HVC tracts in each era.

105. While this first approach is quite intuitive and straightforward, it does not allow me to also control for property crime, accidents, or tract population, which also might affect police deployments. Therefore, I complement this analysis with two other slightly more involved statistical approaches that I discuss in more detail below.

[18] The one exception may be the onset of Covid-19 pandemic, as the shutdown of schools and businesses may have changed the frequency with which people were driving their cars, and hence the frequency with which they could commit traffic-related violations.

b) Multi-Variate Linear Regression Prediction

106. My second approach to assessing racial disparities in citations by neighborhood of issuance follows a two-stage procedure. In the first stage, I use regression methods to estimate how citation activity in Low-Minority tracts is related to a variety of control variables capturing population, recent accidents and recent criminal activity in those tracts. In the second stage, I then use these estimates to predict what citation activity should look like in High-Minority tracts (and Mixed-Race tracts) under the null hypothesis that the relationships between citation activity and these control variables are similar as between High-Minority (and Mixed-Race) tracts, on the one hand, and Low-Minority tracts, on the other. I then assess this null hypothesis by testing whether the actual citations issued in High-Minority tracts (and Mixed-Race tracts) differ from what would be predicted under the null hypothesis. If the null hypothesis is true—i.e., if the relationship between citations and population, recent accidents, and criminal activity is the same in High-Minority (and Mixed-Race tracts) as it is in Low-Minority tracts—then I should find that, on average, there exist minimal differences between actual and predicted citations in High-Minority (and Mixed-Race) tracts. To the extent I find large differences between actual and predicted citations in High-Minority (and Mixed-Race) tracts, this would cause me to reject the null hypothesis, and would indicate that differences in citations between High-Minority (and Mixed-Race) tracts and Low-Minority tracts cannot be explained simply by the difference in accidents, criminal incidents, and populations between these different types of tracts.

107. Given the possible changes in policing practices that occurred over the whole time period I am analyzing (the beginning of 2012 through the end of 2022), I do this analysis separately for each of the five eras I described earlier.

108. In implementing the first stage, I estimate the relationship between monthly citation activity in each Low-Minority tract and population, recent accidents and criminal activity in these tracts using Ordinary Least Squares (OLS) multi-variate linear regression. An OLS regression provides a way to estimate a linear approximation of the Conditional Expectation Function (CEF), which describes how the (mathematical) expectation (i.e., the mean) of some outcome of interest relates to a collection of control variables.

109. Specifically, for each era, using just Low-Minority tracts, I first use OLS to estimate the following regression equation:

$$E[Citations_i] = \alpha + Population_i * \beta_1 + Accidents_i * \beta_2 + Crime_i * \beta_3 \quad (1)$$

where $E[Citations_i]$ is the expected number of citations issued per month in each census tract i during the course of the era, $Population_i$ is the population of tract i, $Accidents_i$ is a vector of the number of minor accidents and injury accidents in each tract i in a time period prior to the start of the era, and $Crime_i$ is a vector containing the number of property crime incidents and violent crime criminal incidents in each tract i in a time period prior to the start of the era. The estimated β coefficients capture how tract differences in the control variables affect the expected number of citations per tract amongst Low-Minority tracts.

110. In the second stage, I then use the estimated coefficients from equation (1) to calculate the expected, or predicted, number of citations in each High-Minority tract (and Mixed-Race) tract given their observed populations, accidents and criminal occurrences. These predictions then allow me to define the following variable for each tract i for each month within each era:

$$Diff_i \equiv Citations_i - (\alpha + Population_i * \beta_1 + Accidents_i * \beta_2 + Crime_i * \beta_3) \quad (2)$$

111. This variable $Diff_i$ captures the difference between the actual number of citations in each tract i and the predicted number of citations in that tract based on the relationship between citations and population, accidents and criminal incidents in Low-Minority tracts. As stated above, if the relationship between population, accidents, crime, and citations is approximately the same in High-Minority (and Mixed-Race) tracts as in Low-Minority tracts, then this variable should on average be near zero across all tract categories. On the other hand, if this variable is significantly different from zero for High-Minority (and Mixed-Race) tracts, this would indicate that differences in population, accidents and/or criminal incidents across tracts cannot fully account for differences in citation activity between High-Minority (and Mixed-Race) tracts and Low-Minority tracts. This variable should on average be close to zero for Low-Minority tracts. That is because the predictions are based on regressions conducted using just the Low-Minority tracts, and so, by construction, the predicted citations and the actual citations should on average be the same in Low-Minority tracts.

112. To formally test this, I estimate the following regression:

$$E[Diff_i] = \rho + \lambda_1 HighMinorityTract_i + \lambda_2 MixedRaceTract_i \quad (3)$$

where $HighMinorityTract_i$ is an indicator for whether tract i is a High-Minority tract, $MixedRaceTract_i$ is an indicator for whether or not tract i is a Mixed-Race tract, and ρ, λ_1, and λ_2 are coefficients indicating the extent to which there is a difference on average between actual and predicted citations for Low-Minority tracts (captured by the intercept ρ), High-Minority tracts (captured by λ_1), and Mixed-Race tracts (captured by λ_2). As stated above, by construction, ρ should be very close to zero. A positive and significant λ_1 would indicate that there are more citations in High-Minority tracts than would be predicted based on their population, prior accidents, and prior criminal incidents, as well as how these variables correlated with citations in Low-Minority tracts (the analogous would be true for positive and significant λ_2 with respect to Mixed-Race tracts).

113. Note that by using the average number of citations per month within each era for each tract as the outcome variable of interest, the results are directly comparable across eras even though eras differ in length. However, given that each era is more than a month long, and I am looking at monthly citation counts for each tract in each era, each census tract contributes more than one observation to the analysis within each era. Since these multiple contributions of each census tract are not likely independent, I account for this by clustering my standard errors by census tract.

114. One decision I must make is the point in time at which to measure my control variables of accidents and criminal activity for these regressions. The methodological difficulty is that changes in police strategy or practices that coincided with the implementation of the Strikeforce, the movement of ticketing management to the BTVA, the dissolution of the Strikeforce, or the onset of the Covid-19 pandemic might not only have affected citations, but also affected accidents and/or crime. For example, if the BPD imposed a change in policing policy at one of these junctures that increased focus on high-accident and/or high-crime areas, accidents or crime may have fallen in those areas. This in turn could have caused those areas to subsequently become lower accident or lower crime areas, but only because of the more active policing in those areas. This might obscure the argument that police policy was to focus policing efforts on areas

that were previously known to have high rates of crime and/or accidents, and thereby could also lead to increases in citations in those areas even if accidents and/or crime were not increasing (or were decreasing) in those areas.

115. To account for this issue, I use lagged values of control variables that reflect the value of these variables prior to key changes in BPD practices. I consider two different approaches for doing this.

116. In the first approach, I measure the number of accidents and criminal incidents in the 12 months prior to the start of the era being analyzed. So, for example, when analyzing the Strikeforce era, which starts in July 2012, I use the number of accidents and criminal incidents in each tract from July 2011 to June 2012. By contrast, when analyzing the BTVA period, which started in July 2015, I use the number of accidents and criminal incidents in each tract from July 2014 to June 2015.

117. The second approach measures accidents and criminal incidents in the 12-month period leading up the creation of the Strikeforce (i.e., July 2011 – June 2012) for all eras. This has the benefit that including these control variables in this way means they will be unaffected by any changes in BPD practices from the creation of the Strikeforce onward. The drawback of measuring these control variables this way is that for the later eras, I am measuring these variables with a considerable lag. I do my analysis both of these ways to again assess the robustness of my results. I include the results from this latter approach in the Appendix (as that Appendix reflects, results are very similar under both approaches).

c) Propensity Score Matching

118. One drawback when using a linear regression methodology to predict out-of-sample as in the methodology outlined above is that if some of the out-of-sample units differ dramatically from almost all of the in-sample units, then one may be relying strongly on the linear specification inherent in an OLS regression. In the context here, if some of the High-Minority tracts had far greater numbers of recent accidents and crime occurrences than *all* Low-Minority tracts, then the estimated coefficients from equation (1) using only Low-Minority tracts would be used to make predictions for tracts with characteristics well out-of-sample. This means that the predicted citations for such High-Minority tracts will rely heavily on linearly interpolating what citation rates would have looked like in Low-Minority tracts if there were such Low-Minority tracts that had the higher number of recent accidents and crimes that occurred in these High-

Minority tracts. Therefore, I also evaluate racial differences citation practices across tracts via Propensity Score Matching. As I discuss in more detail below, this method provides a way to ensure that we are comparing citation counts between High-Minority tracts and Low-Minority tracts that are very similar in terms of recent accidents and criminal incidents.

119. Propensity Score Matching provides an alternative method to multi-variate linear regression for estimating how the conditional mean of some outcome of interest for one group (the *Treated* group) differs relative to another (the *Untreated* group). The basic idea is to first match units of the *Treated* Group with "similar" units of the *Untreated* group, where "similar" is defined with respect to other characteristics that might also affect the outcome of interest. Once these matches are made, the *Treatment* group and these "matched" units from the *Untreated* group should have similar characteristics. We can therefore estimate how the mean outcome for those in the *Treated* group differs from the mean outcome for similar members of the *Untreated* group by simply comparing the average outcomes among the members of the *Treated* group and the *Untreated* group that are "matched" based on their similar characteristics.[19]

120. In the context of this analysis, we can think of whether or not a tract is High-Minority as the "treatment" of interest, and for each era we want to estimate the extent to which average monthly citations differ in High-Minority tracts relative to Low-Minority tracts that are *otherwise similar with respect to accidents and criminal activity*. To do so via propensity score matching, we must start by estimating a propensity score, which in this case is an estimate of the likelihood (or *propensity*) that a given census tract is High-Minority given the prior number of accidents (minor and injury) in the tract and the prior number of serious crime incidents (property and violent) in the tract. This propensity score narrows this multi-dimensional set of tract characteristics into a single-dimensional measure which can be used to directly match High-Minority tracts to Low-Minority tracts with similar propensity scores. This in turn should mean that we are matching High-Minority tracts to Low-Minority tracts that are very similar in terms of the relevant characteristics discussed above, which is something that I can confirm empirically.

[19] Two classic citations for this approach include Rosenbaum, P.R. and D.B. Rubin (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects." Biometrika 70: 41-55, and Heckman, J.J., H. Ichimura, and P. Todd. (1997). "Matching as an Econometric Evaluation Estimator." Review of Economic Studies 65: 261-294.

121. One slight drawback to this approach is that I must limit my sample to only those tracts classified as either High-Minority (i.e., greater than 60% Minority residents) or Low-Minority (less than 40% Minority residents), as this procedure only works for comparing two groups to each other. However, given the focus of this analysis, and given that High-Minority and Low-Minority tracts constitute 70 of the 79 census tracts in Buffalo, omitting the Mixed-Race tracts from this analysis is not overly concerning.

122. To calculate propensity scores for each tract, I use a Probit regression specification to regress an indicator variable for whether or not each tract is High-Minority on the number of minor and injury accidents in each census tract in the year leading up to the creation of the Strikeforce, and the total number of property and violent crimes incidents in each census tract in the year leading up to the creation of the Strikeforce.[20] The reason I use pre-Strikeforce measures of the accident and crime variables is because I want to match the same tracts together in each era. If I measured these control variables over the 12 months prior to the start of each era, which could potentially cause my matched tracts to differ across eras, which may make results across eras not directly comparable.

123. Given the procedure described above, these propensity scores for each tract give the predicted likelihood that a given tract is a High-Minority tract based on the observed characteristics described above. To the extent the tract characteristics described above are positively correlated with whether or not a tract is High-Minority, tracts with higher values for such characteristics will have higher propensity scores. Once a propensity score is estimated for each tract, I then match each High-Minority tract to a Low-Minority tract using *Nearest Neighbor with Replacement*. This method simply matches each High-Minority tract i to the Low-Minority tract whose propensity score is closest to that of tract i.

124. When using this approach, I seek to avoid comparisons of nearest neighbors whose propensity score match isn't "close," as such tracts may be dissimilar to each other with respect to their characteristics. Therefore, for this analysis, I limit my sample to those census tracts within the "common support," or those tracts whose propensity scores fall within the range of propensity

[20] A Probit specification is an alternative estimation procedure to OLS specifically developed for estimating regressions with binary outcome variables. It is particularly important to use in the context of something like propensity scores, as this regression specification will give predicted values for whether or not each observation is treated, where these predicted values are constrained to be between zero and one.

score values that overlap between High-Minority tracts and Low-Minority tracts. This removes any High-Minority tracts that are so different in their accident and/or crime characteristics from all of the Low-Minority tracts such that there are no appropriate matches for them from amongst the Low-Minority tracts.

125. Once I have a match for each High-Minority tract (in the common support), the Propensity Score Matching estimator computes the average difference in monthly citations between each High-Minority tract and its Low-Minority tract match. I again do this separately for each of the five eras described above.

126. I implement this methodology by computing propensity scores and matches using the psmatch2 package in the statistical software program Stata 17. I then take the sample of High-Minority tracts in the common support and their matched Low-Minority tracts, and regress monthly citations within each era on an intercept and an indictor variable for whether the tract is High-Minority. The coefficient on this variable is the test statistic that I report, and which represents the average difference in monthly citation counts between High-Minority tracts in the common support and their matched Low-Minority tracts. I again calculate robust standard errors clustered by census tract to account for non-independence within census tracts over time within an era, and then use a t-test to evaluate statistical significance.[21]

127. As stated above, in performing this analysis, I limit my sample not only to those tracts classified as either High-Minority (greater than 60% Minority residents) or Low-Minority (less than 40% Minority residents), but also only to those tracts whose propensity score lies within the common support. I argue that this makes this methodological approach quite conservative for assessing unexplained disparities in citation counts across tract types because the common support restriction means that I do not include High-Minority tracts that are very different from all Low-Minority tracts. By excluding these tracts from the matching analysis, this analysis focuses exclusively on the High-Minority tracts that had relatively similar prior accident and crime rates to at least some Low-Minority tracts. Therefore, if most of racial inequality in citations across neighborhoods is arising in the subset of High-Minority tracts that do not have characteristics that make them comparable to any Low-Minority tracts, or in tracts that are classified as Mixed-Race relative to Low-Minority, then this approach will actually understate the racial inequality with

[21] These standard errors are also robust to arbitrary forms of heteroscedasticity, or correlation between the size of the variance in citation counts and any of the control variables.

respect to place of citation. Moreover, due to the more limited sample size, the number of observations on which these estimates are based will be lower than in the Regression Prediction methodology discussed above. This generally increases the size of standard errors, making claims of statistically significant differences harder to achieve.

128. To summarize, the Propensity Score Matching estimator tells us the extent to which citation practices differ between High-Minority tracts and Low-Minority tracts that had similar populations and similar numbers of minor accidents, injury accidents, violent crimes, and property crimes in the year prior to the creation of the Strikeforce. This means that this analysis examines only the sample of High-Minority tracts that can be matched to a Low-Minority tract that is quite similar on all of these dimensions. Therefore, the results and the precision of the results (i.e., standard errors and p-values) coming from the Propensity Score Matching analysis could potentially differ from both the analysis of the *High-crime* tracts only and the Linear Regression Prediction analysis described earlier, due to both differences in estimation method and composition of sample. I therefore compare results across these different methodologies to ascertain whether those results are fundamentally consistent or where they may diverge.

2. Methods for Examining Racial Disparities in Citations Across Individuals Within Neighborhoods

129. Above, I described the methods I use to analyze whether citation practices differed across neighborhoods based on the racial composition of those neighborhoods. While these methods address racial disparities with respect to which neighborhoods are cited more intensely, a second issue to analyze is how citation counts relate to the actual racial identity of the cited individuals relative to the racial composition of Buffalo as a whole as well as the racial composition of the neighborhood in which the drivers were cited.

130. For this analysis, I need to know how many tickets were issued to members of each race within each tract in any given time period, which requires information about the race of ticketed drivers. As described in Section IV above, information about the race of cited drivers was contained in two sources: the TraCS citation data and the Open Data Buffalo (OD) dataset that was merged with the TraCS data.

131. From these datasets, I was able to determine whether or not a cited driver was Minority for 82% of TraCS citations using the following procedure. First, for any citation i that had race information directly included in the TraCS dataset, I used the included information about

race.[22] Second, for any citation *i* that did not have race information directly included in the TraCS dataset, but there existed a different observation *j* with the same DL number as observation *i* with the same DL number and had race information directly included in the TraCS dataset, my TraCS race measure for *i* corresponds to the TraCS race of observation *j*. Third, for any observation *i* still missing race, but for which there existed another observation *j* within the same TraCS *incident* (as defined in the Data section above), my race measure for observation *i* corresponds to the TraCS race of observation *j*. In a very small number of cases, the race of driver with same DL number is not always the same. This likely arises due to coding inconsistencies by the BPD. In these cases, I assigned the modal race associated with that DL number.

132. Fourth, for the remaining citations that still did not have a recorded race, I then turned to the OD data.[23] In the OD data, I first defined an incident, which I identified using the procedure described above in the Data section. Within each incident, I then took the modal race and modal ethnicity of all the observations in the incident (these were recorded separately in OD). In most cases, all observations within the same incident had the same recorded race and ethnicity, or were a combination of observations with the same race/ethnicity and observations missing race/ethnicity information. However, for a very small number of incidents, there were observations of conflicting race (12) or conflicting ethnicity (4). For these observations, I counted that OD variable as missing. I combine the race and ethnicity information into a binary variable equaling one if Minority (i.e., Black and/or Hispanic) or zero otherwise, which I again refer to as "race".

133. Once I determined race for OD incidents (when possible) using the above procedure, I then matched these OD incidents to TraCS incidents using the procedure laid out above in the Data section. For any observation that had no TraCS race recorded, but which could

[22] Note that race and ethnicity were recorded using the same variable in TraCS, so I refer to race/ethnicity simply as "race" here. In reality, "race" and "ethnicity" refer to different constructs that are not mutually exclusive. For example, an individual may be both White and Hispanic or Black and Hispanic. *See, e.g.*, U.S. Census Bureau, *Hispanic Origin*, https://www.census.gov/topics/population/hispanic-origin.html (revised Apr. 26, 2024). Because I have treated all Hispanic individuals as "Minority," this distinction does not affect my results.

[23] The reason I used the TraCS race rather than OD race when both were available for a given observation was because my understanding is that the race in TraCS corresponds to what the officer issuing the ticket inputted at the time of issuing the ticket. By contrast, OD race may be linked to a given citation based on the race input by an officer in a previous encounter that person had with police. As I show later, however, for citations with race data from both TraCS and OD, race is almost always the same.

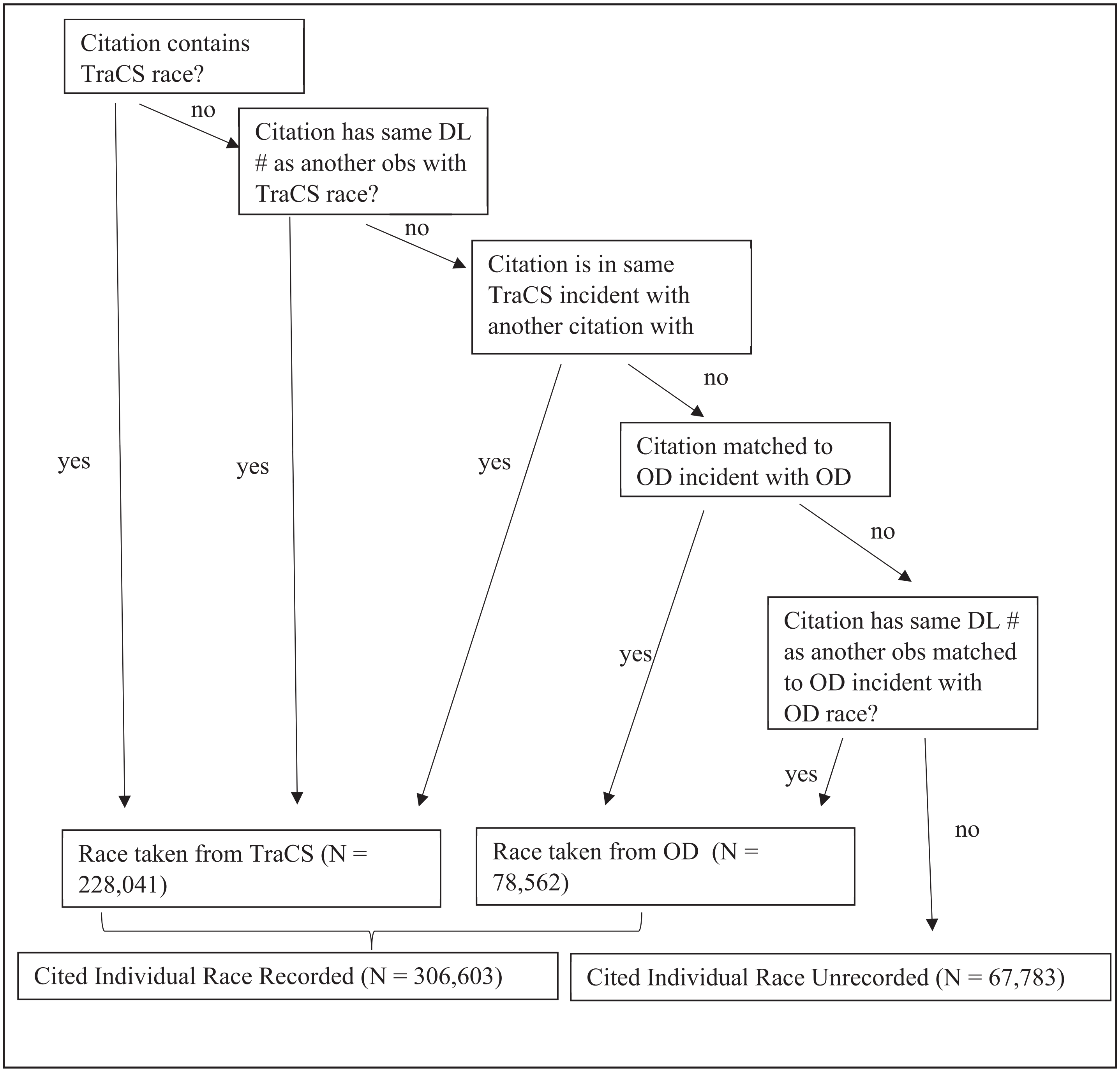
be matched to an OD incident with valid OD race, my race measure corresponds to the OD measure. Finally, for any observation with no recorded TraCS race that was not matched to an OD incident with valid OD race, but which had the same DL number as another observation that was matched to an OD incident with valid OD race, my race measure corresponds to that valid OD race measure. Where race data could not be obtained directly from TraCS or OD, I consider race to be unrecorded.

134. In summary, for any citation *i* with TraCS race information, my race measure corresponds to that TraCS race. For any observation *i* that did not have TraCS race information, but did have OD race information, my race measure for observation *i* corresponds to OD race. Any observation where I could not determine race from either source, I count has having race unrecorded (about 18% of the total observations).

135. The reason I use the TraCS race information as my primary source for driver race is because my understanding is that this comes from the officer at the time the citation was issued. My understanding is that in the OD, the citation was linked to the broader CHARMS system, which may have had race information that was gathered from a previous incident, in which a different officer made the race categorization.[24] Given my analysis focuses on the citation at hand, it seems like the TraCS categorization is the more relevant one. However, as I discuss later in the paper, when both TraCS and OD contain race information for a given incident, they are consistent over 96% of the time.

136. The following figure describes graphically how I used information from both of these data sources to assign race to observations.

[24] Email from Robert Quinn, Corporate Counsel for the City of Buffalo, dated December 17, 2021.



137. Given this measure of race of the cited individual, I begin by performing the following descriptive analysis. For each of the five eras described previously, I compare the mean number of monthly citations issued citywide to Minority drivers, Non-Minority drivers, and drivers of unrecorded race. Notably, I also do this separately for tracts classified as High-Minority, tracts classified as Low-Minority, and tracts classified as Mixed-Race. This gives us a broader

picture of the racial composition of ticketed drivers in the different types of neighborhoods over time.

138. As a more formal approach to assessing racial disparities in citation activity within neighborhoods, I again use census tracts as my definitions of neighborhoods and compute the following Population Adjusted Ticket Rate *(Pop Adj Tix* $Rt_{r,t}$*)* for each tract t for each race r for each era:

$$Pop\ Adj\ Tix\ Rt_{r,t} = \frac{\sum_{m=1}^{M} tickets\ issued_{r,t,m}/M}{population_{r,t}/100}, \qquad (4)$$

where $tickets\ issued_{r,t,m}$ refers to the number of tickets issued to individuals of race r in tract t for each month m within any era, M is the number of months in the era, and $population_{r,t}/100$ refers to the number of individuals (in hundreds) of race r living in tract t. In words, *Pop Adj Tix* $Rt_{r,t}$ gives the average number of tickets issued per month per hundred population for each race r and each census tract t in a given era.

139. To the extent tickets are issued in proportion to the population of each race within each census tract, then on average across tracts, *Pop Adj Tix* $Rt_{r,t}$ should be roughly even across racial groups (i.e., Minority and Non-Minority) within each tract. However, since tracts differ in size and the goal is to test whether *Pop Adj Tix* $Rt_{r,t}$ differs on average across the Buffalo population, I compute the weighted average of (4) across all tracts for each race r, where I weight by total tract population, so that the weighted average will capture the *Populations Adjusted Ticketing Rate* for the average citizen, which will allow for direct comparisons to citywide *Population Adjusted Ticketing Rates*.

140. Once I compute *Population Adjusted Ticketing Rates* for each tract/race in each era, I use a t-test to test the null hypothesis that the population weighted average of *Pop Adj Tix* $Rt_{r,t}$ across tracts is the same across races within each era. I again cluster my standard errors by census tract when computing t-statistics and determining statistical significance. If these statistics differ significantly across races within any given era, this would imply that, on average, there is a racially disproportionate number of tickets issued relative to population within tracts.

141. One important issue that arises in calculating the *Population Adjusted Ticketing Rate* for each race in each tract as defined by equation (4) is that race is not recorded for just under

20% of observations. As a result, there is the possibility that if the distribution of race differs among these observations for which race is unrecorded relative to those observations for which race is recorded, then assessing whether the average tract level *Population Adjusted Ticketing Rate* differs across races may be statistically biased. Therefore, in order to assure that the average *Population Adjusted Ticketing Rate* for each race is statistically unbiased, I must determine a reasonable way to assign the tickets issued to drivers for whom race is unrecorded.

142. To do so, I divide all citations into 750 different possible cells based on the racial composition type of each driver's zip code of residence (0-20% Minority, 20-40% Minority, 40-60% Minority, 60-80% Minority, 80-100% Minority), the racial composition type of the tract of issuance (0-20% Minority, 20-40% Minority, 40-60% Minority, 60-80% Minority, 80-100% Minority), whether the ticketed individual's zip code of residence was in Buffalo/first ring suburb of Buffalo/Outside of Buffalo area, whether the citation is a Moving violation or a Non-Moving violation, and which of the five eras the citation was issued. I then use just those citations for which information on race was recorded in the TraCS data to compute the average fraction Minority among individuals in each of these cells.

143. Under the assumption that it was effectively random whether or not race was recorded in TraCS among drivers residing in zip codes with similar racial demographics, who were stopped in census tracts with similar racial demographics, who were cited for similar infractions, who were from the same general region around Buffalo, and who were cited in the same time era, the racial composition of those with unrecorded race should be similar to the racial composition of those with recorded race in TraCS *within* each of the 750 possible cells described above. Therefore, for each cell, I can impute the fraction of citations of unrecorded race that were Minorities versus Non-Minorities based on the fraction Minority among those with TraCS race in that cell.

144. Based on these imputations for each cell, I can then aggregate up the imputed fraction Minority amongst all of the citations with unrecorded race issued in each census tract in each era, and thereby can estimate of how many citations were issued overall to Minorities and Non-Minorities in each census tract in each era. This allows me to calculate measures of the *Population Adjusted Ticketing Rate* for each race for each tract across eras and perform the t-test described above.

145. To illustrate this procedure with a simplified example, suppose that in a given era in a given tract, there were 200 citations with unrecorded race, with 120 coming from some cell #1 and the remaining 80 coming from some cell #2 (where these cells were determined using the driver and citation characteristics discussed above). In cell #1, the average fraction Minority among those with race recorded in TraCS was 80%, while in cell #2, the average fraction Minority among those with race recorded in TraCS was 25%. This would mean that 0.8*120 + 0.25*80 = 116 of the citations with unrecorded race in this tract in this era would be counted as going to Minorities under the above procedure, with the remaining 84 counted as going to Non-Minorities. If there were another 600 citations in this tract in this era with race recorded (either in TraCS or OD), with 400 being issued to Minorities and 200 issued to non-Minorities, then overall in this tract in this era, 400+116= 516 would be counted as issued to Minorities and the remaining 200+84= 284 would be counted as issued to non-Minorities. These would be the citation counts used in equation (4) for each race for this tract in this era.

146. As I show in the Analysis section below, this imputation methodology has very high internal validity in the sense that when I use this procedure to estimate the allocations of citations across races amongst those with recorded race, it almost perfectly matches the true allocations of citations across races amongst this population. Moreover, as also shown below, this procedure suggests that the fraction Minority amongst those with unrecorded race is somewhat lower than it is among those with recorded race (in other words, the BPD officers were a bit less likely to record race when citing Non-Minority individuals). This means failing to include those with unrecorded race in computing *Population Adjusted Ticketing Rates* would lead to statistically biased results, in the sense that computing Population Adjusted Ticketing Rates without including the citations with unrecorded race would generally overstate the true difference in such rates between Minorities and Non-Minorities. Therefore, the results I show below include those with unrecorded race via the imputation procedure laid out above.

147. As alluded to above, I can also aggregate further up to the citywide-level, and estimate how many overall tickets were issued to Minorities citywide in any given era, as well as how many overall tickets were issued to Non-Minorities citywide in any given era. This allows me to compute a city-wide version of *Population Adjusted Ticketing Rates* for each race, rather than just tract level rates. As I discuss below, given my average tract level *Population Adjusted Ticketing Rates* are weighted by tract population, I can use my average tract level *Population*

Adjusted Ticketing Rates for each race to decompose racial differences in the overall citywide *Population Adjusted Ticketing Rates* into the portion due to within-tract disparities and the portion due to across-tract disparities.

3. Methods For Analyzing Racial Disparities Across Incidents

148. Another issue for analysis is whether there are racial differences in treatment conditional on being stopped. In particular, I again use individual-level race data to analyze whether there are racial differences in the likelihood of receiving multiple tickets within the same incident, where an incident includes all the citations issued at a particular location on the same date, at the same time (to the minute), and to the same person. If there is more than one citation issued within an incident, I count that as an incident that led to multiple citations.

149. To analyze racial differences in the likelihood of receiving multiple tickets within the same incident, I use a two-stage procedure that is analogous to the one I used for analyzing citations by neighborhood, except that it uses incidents rather than citations as the unit of analysis. Namely, for each of the five eras, I first use OLS to estimate the following regression using only the sample of incidents issued to Non-Minorities:

$$E[Multipletix_i] = \alpha + Age_i * \beta_1 + Time_i * \beta_2, \quad (7)$$

where E[$Multipletix_i$] is the expected likelihood incident i resulted in multiple citations, Age_i is the age of the driver, and $Time_i$ is a vector of indicator variables associated with the time of day of the incident. I control for age of driver and time of day of the stop in case (i) BPD practices with respect to multiple ticketing are correlated with these incident characteristics, and (ii) these characteristics are correlated with driver race. The β terms refer the estimated coefficients that capture how the likelihood that an incident resulted in multiple citations is affected by age of driver and by when the incident occurred in incidents involving Non-Minorities.

150. In the second stage, I use the estimated parameters from equation (7) to predict the likelihood that each incident involving a Minority driver leads to multiple citations (as well as the likelihood that each incident involving someone of unrecorded race leads to multiple citations), and compute a new variable capturing the difference between the actual incidence of multiple citations to this predicted likelihood for each of these incidents. I then regress this difference

variable on an indicator for whether or not the incident involves a Minority and an indicator for whether or not the incident involves an individual of unrecorded race (along with an intercept term, which gives the average difference between actual and predicted multiple citations for incidents involving Non-Minorities, and therefore should be very close to zero by construction). If the coefficient on the indicator for the incident involving a Minority is positive, this would imply incidents involving Minorities are more likely to lead to multiple citations relative to incidents involving Non-Minorities that occurred during the same time of day and drivers of the same age (similar interpretation for the coefficient on the indicator for unrecorded race).

151. To the extent that this method shows that Minority individuals are more likely to receive multiple tickets per incident than Non-Minority individuals, this could possibly result from either racial bias by the BPD or because stopped Minority individuals may have been more likely to commit multiple infractions than Non-Minority individuals. Therefore, I also examine a more specific type of multiple ticketing that is less subject to such uncertainty. In particular, I assess whether there is a racial difference in the likelihood of receiving multiple tinted window tickets within incidents involving a tinted window ticket. I am not aware of any evidence suggesting that among drivers who tint their windows, Non-Minorities are more likely than Minorities to tint only one window (indeed, it would seem very odd for a driver of any race to tint only one window on a car). Therefore, to the extent that there are racial differences in the likelihood of receiving multiple tinted window tickets within tinted window incidents, it is difficult to think of any race-neutral explanation for that discrepancy.

152. To analyze multiple tinted-windows ticketing, I perform analogous analyses to those described above, but limit my sample to just incidents that involve tinted windows and focus only on whether or not there are multiple tinted window citations within such incidents.

C. Methods for Examining Racial Disparities in Stops.

153. I understand that, starting in 2020, the BPD agreed to document traffic stops that led to a stop receipt but no actual citation.[25] The data do not indicate why these stops did not lead to citations (for example, where closer inspection revealed that there was no violation versus where the officer believed that there was a violation but simply chose not to issue a citation).

[25] Mayor Byron W. Brown, Executive Order No. 2020-001, https://www.buffalony.gov/DocumentCenter/View/7602/OA-ExecutiveOrder-Policefinal.

154. To analyze these stop receipts, I focus on how the reasons for such stops differed across BPD districts, which differed substantially in their racial demographics, as well as how the reasons for such stops differed by race of driver. The one constraint with the latter is that driver race is missing for approximately 20% of recorded stops receipts. Moreover, the stop receipts data lack information about the driver's zip code of residence for the majority of observations, so I cannot do a procedure analogous to what I did with respect to citations to impute the racial composition of those for whom race is unrecorded. Therefore, I consider three categories of driver (i) Minority, (ii) Non-Minority, and (iii) Race Unrecorded.

155. To perform my analysis, I first classify all stops into one of 12 relatively broad categories (equipment, failure to stop at stop sign or stop light, registration/insurance violation, speeding, failure to use turn signal, erratic/reckless/illegal driving, investigation/stolen vehicle/warrant, suspicion of crime/suspected of crime/drugs, illegal parking/stoppage, cell phone/texting, seatbelt/child-seat violation, or no reason given). From these, I then focus on the first 6 of the above categories, which were the categories that contained the most observations, encompassing almost 94% of all stop receipts.

156. I then examine how the fraction of stop receipts in each of these categories differs between the BPD districts with the highest Minority populations and the BPD districts with the lowest Minority populations. The reason I use the BPD district as opposed to census tract (as I did with citations) to locate stop receipts is because 32% of stop receipts were not able to be mapped to census tracts in the data I was given. By contrast, the BPD recorded the BPD police district of the stop for all stops in this dataset.

157. I also then examine how the fraction of stop receipts in each of these categories differs between Minority drivers and Non-Minority drivers. As discussed earlier, race of driver is missing for about 20% of stops receipts. Moreover, the zip code of residence was missing for a large number of drivers in the stops receipt data, making me unable to use the imputation method for allocating observations with unrecorded race as I did with the citations data. Therefore, in this analysis, I simply use "race unrecorded" as a distinct category.

D. Methods for Examining Racial Disparities in Checkpoint Locations

158. I was given data on the locations of the over 1,600 checkpoints the BPD set up in the City of Buffalo between January 1, 2013, and October 27, 2017. My aim is to analyze whether BPD chose to locate these checkpoints disproportionately in Minority neighborhoods.

159. Given that all of these checkpoints occurred after the implementation of the Strikeforce in mid-2012 and before the end of 2017, the eras I describe above with respect to the analysis of citations are not relevant (i.e., they all took place during the Strikeforce or BTVA eras described earlier). However, one consideration is that in July 2017, the Buffalo Common Council adopted a resolution that raised concerns about discriminatory practices used by the BPD with respect to the use and location of checkpoints. Therefore, in my analysis of checkpoints, I not only consider the whole time period during which checkpoints were conducted according to the data provided by defendants (from January 2013 to October 2017), but also focus on the period preceding the Common Council resolution (January 2013 to June 2017).

160. I first just examine the relative counts of checkpoints across the different categories of census tracts (i.e., High-Minority, Low-Minority, and Mixed-Race). However, in conducting my analysis, I also consider possible explanations for why the BPD located checkpoints where they did that may simply be correlated with the racial composition of the neighborhood. For example, analogous to the issues that I discussed above with respect to citations, the BPD may have chosen to locate checkpoints in areas that had seen a lot of crime in the recent past in an effort to deter subsequent crime.[26] If such locations also happened to often lie in High-Minority neighborhoods, this could potentially account for any empirical relationship between checkpoint location and racial make-up of the neighborhood.

161. Given the similarities in the issues that arise in this context as with respect to the previous analysis of incidence of citations by location, I take a similar approach to analyzing the location of checkpoints as I did with respect to the incidence of citations by location. Namely, I first examine checkpoint counts in only the sample of *High Violent Crime (HVC)* tracts (as defined previously), and test whether there is a significant difference in checkpoints in High-Minority *HVC* tracts relative to Low-Minority *HVC* tracts.

162. Second, I employ a two-stage regression prediction procedure, in which I use OLS regressions to predict the number of checkpoints in each census tract based on the prior criminal incidents as well as prior accidents in that tract and how such variables correlated to checkpoints

[26] I understand that Defendants stated that the checkpoints were not primarily intended to deter crime but may have had "high police visibility" as a "secondary purpose." Defendants' Amended Responses to Plaintiffs' 5th Set of Interrogatories ¶ 19.

in Low-Minority tracts. I then calculate the difference between the actual number of checkpoints and the predicted number of checkpoints for each tract, and assess to what extent this difference differs from zero for High-Minority (and Mixed-Race) tracts.

163. Third and finally, I use Propensity Score Matching to find High-Minority tracts for which there are Low-Minority tracts that had similar numbers of prior accidents and criminal incidents, and then directly compare the number checkpoints that occurred in these High-Minority tracts relative to the matched Low-Minority tracts. The basic procedure for this was described in detail in the sub-section above regarding methods for analyzing racial disparities in citations by neighborhood of issuance. Again, in doing this Propensity Score Matching analysis, I limit my sample to only High-Minority tracts that lie in the common support.

164. Finally, while my main analysis of checkpoints employs the original list of checkpoints provided by the City of Buffalo, counsel has provided me a list of checkpoints that are contained in the original list but which appear not to have actually been implemented based on a review of citations data. I assess the robustness of my results to deleting the checkpoints on this auxiliary list.

VI. ANALYSIS

165. I now present the results of my analysis using the methodologies discussed above. My principal conclusions are stated in the Summary of Findings above. This section describes those conclusions in greater detail.

A. Analysis of Citation Activity Over Time

166. My analysis reveals that BPD ticketing policies and/or practices changed significantly between each of the defined time eras that I defined earlier. Notably, monthly citation rates significantly increased with the implementation of the Strikeforce, and increased again when the BTVA took over ticket adjudication from the State of New York. Monthly ticketing then fell when the Strikeforce was disbanded, and then fell again with the onset of the COVID-19 pandemic. I discuss the details of these findings below.

167. As discussed in the Methods section above, I use Wilcoxon Rank-Sum tests to evaluate whether there are statistical differences in monthly citation volume across adjoining eras, where these eras correspond with the key dates delineated earlier: the Pre-Strikeforce era (January 2012–June 2012), the Strikeforce era (July 2012–June 2015), the BTVA era (July 2015–January 2018), the Post-Strikeforce era (February 2018–February 2020), and the Current era (March 2020–

December 2022). For these tests, I use all observations in the TraCS dataset during these eras. Table 1 presents the results of these Wilcoxon Rank-Sum tests. The first column shows the two eras being compared; the second column shows the number of monthly observations in each era; the third column shows the actual sum of the ranks of each month in the combined sample for each era; and the fourth column shows the expected sum of the ranks in the combined sample for each era under the null hypothesis that the same underlying citation issuance practices were occurring in both eras.

Table 1 - Wilcoxon Rank-Sum Tests of Equality of Distributions Across Eras

		Rank Sum	
Compared Eras	**Obs**	**Actual**	**Expected Under Null**
Era 1 - Pre-Strikeforce	6	24	129
Era 2 - Strikeforce	36	879	774
			z = -3.774
			pval < 0.0001
Era 2 - Strikeforce	36	803	1224
Era 3- BTVA	31	1475	1054
			z = -5.294
			pval < 0.0001
Era 3- BTVA	31	1221	883.5
Era 4 - Post-Strikeforce	25	375	712.5
			z = 5.563
			pval < 0.0001
Era 4 - Post-Strikeforce	25	1156	750
Era 5 - Current	34	614	1020
			z = 6.228
			pval < 0.0001

168. The z is the test statistic of the null hypothesis and has an approximately normal distribution centered at zero with a standard deviation of one. Values further from zero imply that it is less and less likely that the monthly citation counts in the two successive eras come from the same underlying data generating process, which in this case would correspond to the BPD ticketing policies and practices. Negative z values mean there are more citations in the latter era than

expected if monthly citations in both eras come from same underlying process. Positive z values imply that there are fewer citations in the latter era than expected if citations in both eras come from same underlying process.

169. The pval gives the p-value associated with the z statistic under the null hypothesis or the probability that the z-statistic can be so far from zero if indeed the monthly citations in each era come out of the same underlying practices.

170. As can be seen in of Table 1, the Wilcoxon Rank-Sum test strongly rejects each of the null hypotheses that monthly citations in any two successive eras come from the same underlying stochastic process, all with p-values less than 0.0001. In other words, this evidence strongly suggests that BPD ticketing practices changed between each of these defined time eras.

171. As discussed in the Methodology section, to address the possibility that the Wilcoxon Rank Sum test could be overly sensitive, for each era I randomly bifurcated months within the era into two groups and then did Wilcoxon Rank-Sum tests on these two within-era groups. I did ten different random bifurcations for each era, and therefore ten different Wilcoxon Rank-sum tests for each era. To the extent citations within each era come from the same underlying stochastic processes across months within that era, then such tests should rarely reject the null hypothesis that the samples from the two random groups delineated within each era come from the same underlying random process.

172. For each era, Table 2 shows the number of times the pval on such Wilcoxon Rank-Sum tests was below 0.01, the number of times the pval of the on such Wilcoxon Rank-Sum tests ranges between 0.01 and 0.05, the number of times the pval ranges between 0.05 and 0.10, and the number of times the pval of the on such Wilcoxon Rank-Sum tests exceed 0.10. As can be seen, for these within era tests, pvals never fall below 0.01, only twice fall between 0.01 and 0.05, and only four times do such pvals fall between 0.05 and 0.10. In the remaining 44 out of the 50 within era tests these pvals are over 0.10.

Table 2 - Randomized Wilcoxon Rank-Sum Tests Within Eras (10 trials per era)				
	P-values on Null Hypothesis:			
Era	**< 0.01**	**0.01 - 0.05**	**0.05 - 0.10**	**> 0.10**
Era 1 - Pre-Strikeforce	0	0	0	10
Era 2 - Strikeforce	0	1	1	8
Era 3 - BTVA	0	0	0	10
Era 4 - Post-Strikeforce	0	1	1	8
Era 5 - Current	0	0	2	8

*Each trial randomly divides months in each era into two distinct samples. P-values reveal the likelihood these samples come from the same underlying data generating process.

173. It is not surprising that there are a handful of p-values that fall below 0.10 in Table 2 (generally the upper end of what is considered statistically significant), as a p-value in this context gives the likelihood that a given z statistic would arise even if the null hypothesis of no difference is true. Hence, in a sample of 50 tests as we have here, we should expect roughly five would be significant at the 10% level even in this placebo test, which is consistent with Table 2.

174. In summary, Tables 1 and 2 show that the Wilcoxon Rank-Sum tests rejects the null hypothesis of no difference *across* each consecutive eras, while rejection of the null hypothesis of no significant differences in citation practices *within* each era is quite rare. This exercise suggests that the Wilcoxon Rank-Sum test is not prone to finding significant differences between two samples of monthly citations. The eras defined above thus appear to represent significantly distinct time periods with respect to the underlying BPD citation practices.

175. This analysis cannot directly identify why ticketing volume changed across these different eras. However, nothing that I know of would cause Buffalo citizens to commit more motor vehicle code violations following the implementation of the Strikeforce in June 2012 or after the BTVA take-over of ticketing and revenues in July 2015, which would suggest the increase in ticketing volume after these dates reflected a change in BPD policy and/or practices. Similarly, nothing I know of would cause Buffalo citizens to decrease their rate of motor vehicle code violations following the disbanding of the Strikeforce in early 2018, again suggesting that the decline in the number of citations issued after this date reflects a change in BPD policy and/or practices. The exception to this would be the arrival of the COVID-19 pandemic in March 2020, which caused large and abrupt changes in people's commuting behavior and could have led to fewer motor vehicle infractions being committed in the time after the onset of the pandemic.

However, despite the lifting of pandemic restrictions by mid-2021, citations remained lower than pre-2020 levels after mid-2021.

B. Analysis of Racial Disparities in Citations

176. This section presents my analysis of racial disparities in citations. I first do my analysis by neighborhood of issuance, followed by my analysis by race of cited individual.

1. Analysis of Racial Disparities in Citations Across Neighborhoods

177. My analysis reveals that following the implementation of the Strikeforce, a large disparity in monthly ticketing rates arose between neighborhoods with high Minority populations and neighborhoods with low Minority populations. The ratio of monthly citation counts in High-Minority neighborhoods to citation counts in Low-Minority neighborhoods increased from approximately parity in the months before implementation of the Strikeforce to 2–3 in the years after the Strikeforce was implemented.

178. This disparity in monthly ticketing counts between High-Minority and Low-Minority neighborhoods narrowed but remained significant following the disbanding of the Strikeforce. After the onset of the COVID-19 pandemic, the disparity in monthly ticketing counts between High-Minority and Low-Minority neighborhoods mostly dissipated.

179. The large disparities in monthly ticketing counts between High-Minority and Low-Minority neighborhoods between the implementation of the Strikeforce and the onset of the COVID-19 pandemic arose among both Moving violations and Non-Moving violations, though the disparity was approximately 10 times larger for Non-Moving violations than Moving violations.

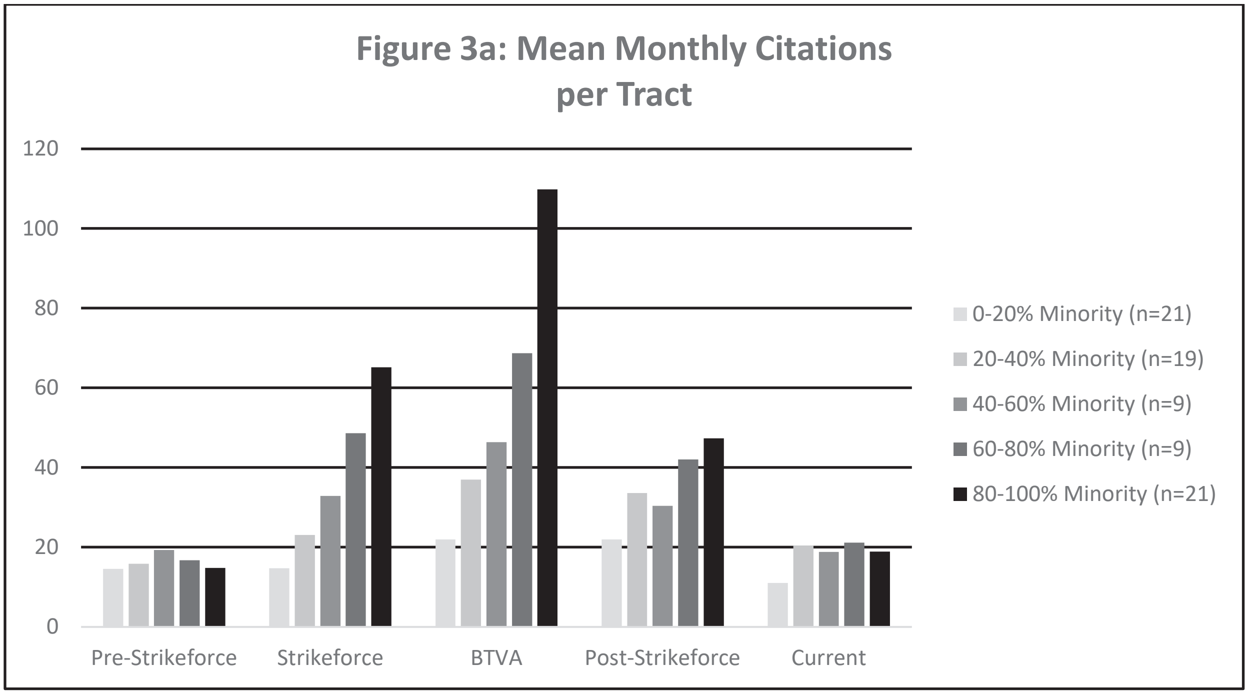
180. Finally, I find that these disparities in Non-Moving violation citations between high Minority neighborhoods and low Minority neighborhoods cannot be explained by the BPD simply focusing their resources on neighborhoods with high crime or a high number of accidents.

181. I discuss the details of my analysis below.

a) Overall Racial Disparities in Citations Across Neighborhoods

182. As a first look at citation patterns in higher Minority versus lower Minority neighborhoods, I examined the mean monthly citations per census tract. Figure 3a shows the mean number of monthly citations issued per census tract, where tracts are grouped into the five categories based on their racial make-up: 0-20% Minority (n=21), 20-40% Minority (n=19), 40-

60% Minority (n=9), 60-80% Minority (n=9), and 80-100% Minority (n=21), for the five time eras described earlier.



183. As can be seen in the left-most set of bars in Figure 3a, in the Pre-Strikeforce era (January 2012–June 2012), the mean monthly citation counts per tract were essentially the same across the five tract racial demographic categories, at a little under 20 per month.

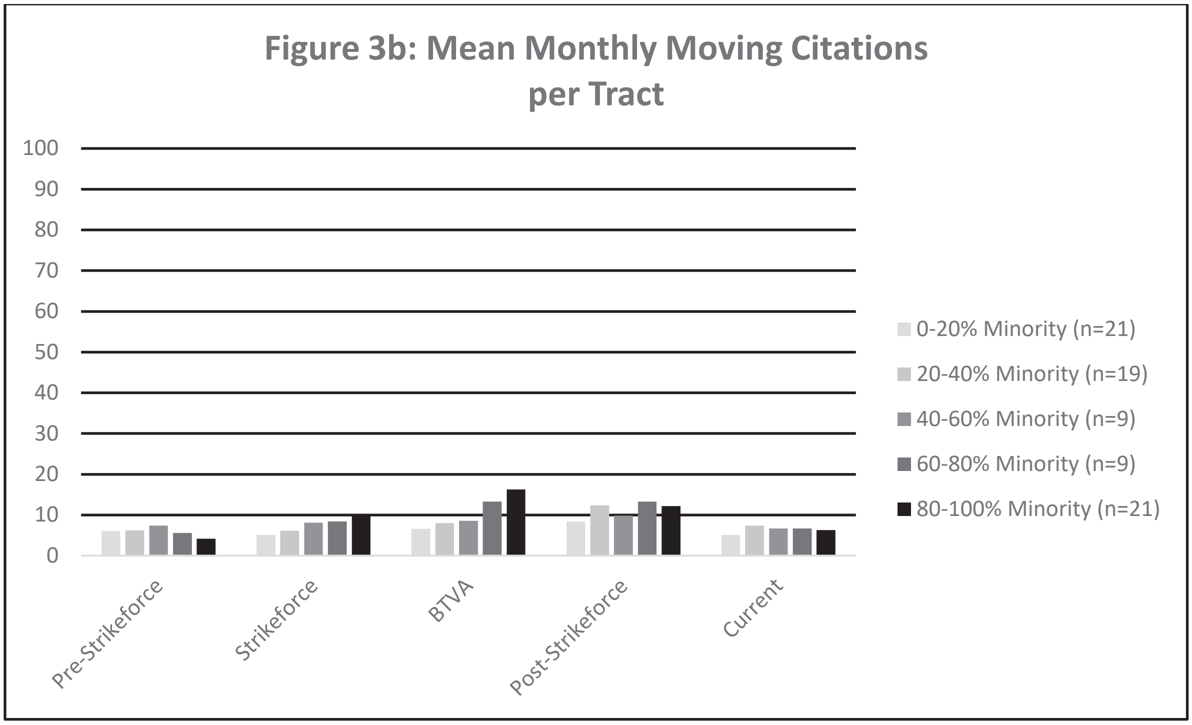
184. The next set of bars in Figure 3a shows mean monthly citation counts per tract in the Strikeforce era (July 2012–June 2015) and reveals that the mean number of monthly citations in the 0-20% Minority tracts stayed roughly constant across the Pre-Strikeforce and Strikeforce eras, and increased modestly in the 20-40% Minority tracts. By contrast, mean monthly citations essentially tripled and more than quadrupled between the Pre-Strikeforce and Strikeforce eras in the 60-80% Minority and 80-100% Minority tracts respectively.

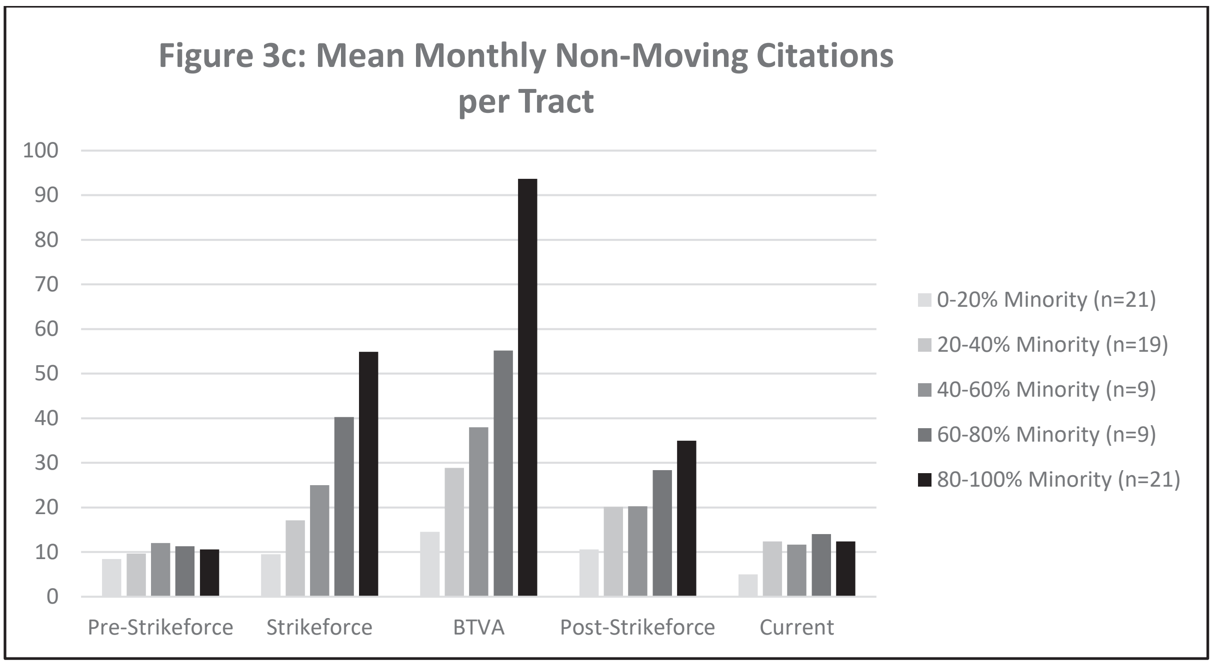
185. Moving to the BTVA era (July 2015–January 2018) bars in Figure 3a, we see that disparities in average monthly citations between tracts with relatively low Minority populations and tracts with relatively high Minority populations increased even more in this era. The lower-Minority tracts averaged under 40 citations per month during this era, while the 60-80% Minority tracts averaged around 70 citations per month and the 80-100% Minority tracts averaged around 110 citations per month during this era.

186. Looking at the bars corresponding to the Post-Strikeforce era (February 2018–February 2020) in Figure 3a, we see that the disparities in mean monthly citations by tract category drop relative to the two earlier eras. This is almost all due to a fall in monthly citations in tracts that were greater than 40% Minority. However, it is still the case that, on average, there were about twice as many monthly citations in the 80-100% Minority tracts relative to the 0-20% Minority tracts during this era. Again, this is in contrast to essentially equal numbers of monthly citations across these different types of tracts in the Pre-Strikeforce era.

187. Finally, looking at the right-most set of bars corresponding to the Current era in Figure 3a (March 2020–December 2022), mean monthly citations fall in all types of tracts relative to the previous era, but again most steeply in the tracts with higher Minority populations. Citations in the Current era more closely resemble citations in the Pre-Strikeforce era, with around 20 citations issued per month across the different tract categories, except for the tracts with the smallest Minority populations (0-20%), which averaged notably fewer citations (just over 10 per month) than tracts with higher Minority populations.

188. I also analyzed the same set of citations by considering whether they were issued for Moving violations (which include violations speeding, reckless driving, failure to stop, DUIs, and improper use of mobile devices) versus Non-Moving violations (which include equipment violations, license/registration violations, insurance violations, and inspection violations). Figures 3b and 3c show how mean monthly citations differed across tract types over these two different categories of citations.





189. As can be seen in Figure 3b, disparities in citations for Moving violations are relatively small across tract categories across four of the five eras. The exception is the BTVA era, where there were on average about twice as many Moving violation citations per month in the 60-80% Minority and 80-100% Minority tracts relative to the 0-20% Minority and 20-40% Minority tracts.

190. As can be seen in Figure 3c, disparities between higher and lower Minority tracts with respect to citations for Non-Moving violations are much more substantial in the Strikeforce, BTVA, and even Post-Strikeforce eras. Once again, the left side of Figure 3c shows that in the Pre-Strikeforce era, average monthly citations for Non-Moving violations are almost identical for the higher Minority tracts as the lower Minority tracts, at about 10 per month. However, moving to the Strikeforce era, average monthly citations for Non-Moving violations jump to 40 and 55 per month in the 60-80% Minority and 80-100% Minority tracts, respectively—a more than fourfold increase for each of these tract categories relative to the Pre-Strikeforce era. This compares to no increase in average Non-Moving citations per month in the 0-20% Minority tracts, and an increase to an average of just 17 Non-Moving citations per month in the 20-40% Minority tracts.

191. Looking at the bars corresponding to the BTVA era in Figure 3c, we see that there are two to four times the number of citations for Non-Moving violations in the 60-80% Minority and 80-100% Minority tracts relative to the lower-Minority tracts in this era. If one looks at just the 80-100% Minority tracts, we see that on average, almost 100 Non-Moving violation citations were issued per month during this era in each of these tracts. This compares to an average of less than 15 such citations per month in the 0-20% Minority tracts during this era. This is again notable given the number of such Non-Moving citations per month were essentially equivalent between these two tract types during the Pre-Strikeforce era.

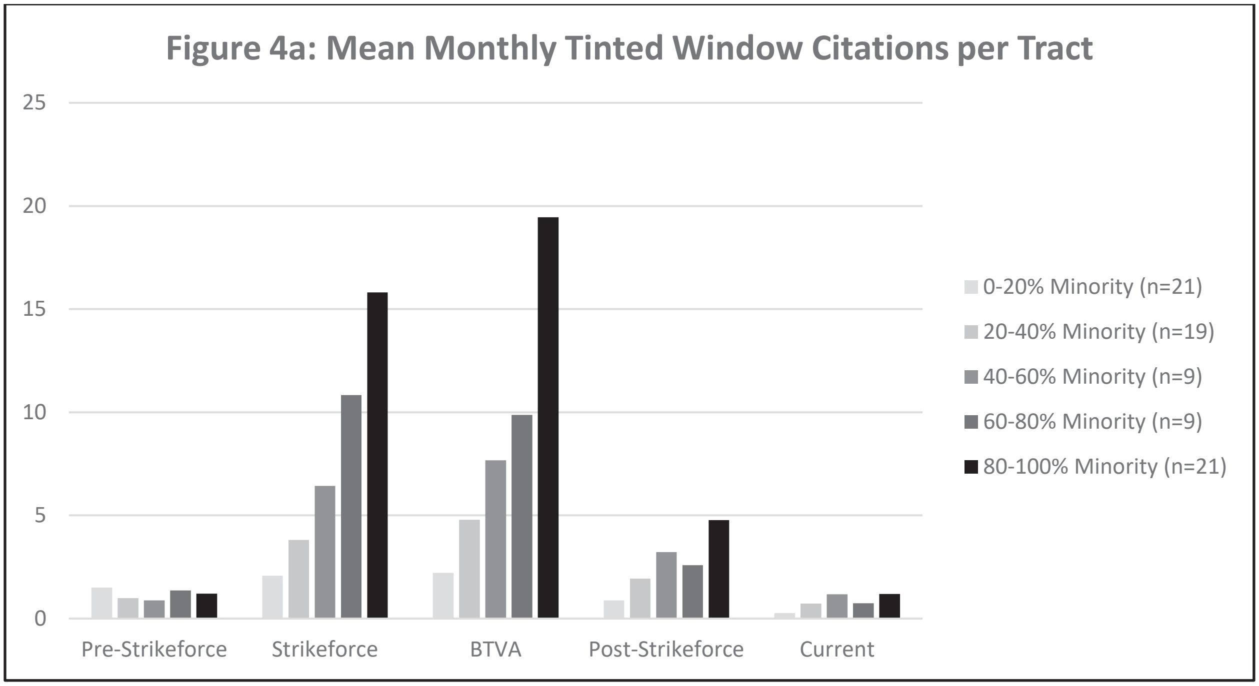
192. Looking at the fourth set of bars in Figure 3c, we see that the disparities in citations for Non-Moving violations between the higher- and lower-minority tracts in the Post-Strikeforce era diminish substantially relative to the BTVA and Strikeforce eras but are still notable. For example, average monthly Non-Moving violation citation counts in the 80-100% Minority tracts were still almost twice as high as what occurred in the 20-40% Minority tracts and three and half times higher than what occurred in the 0-20% Minority tracts during this era.

193. The fifth set of bars in Figure 3c show that during the Current era, there are only small discrepancies between the higher and lower Minority tracts in terms of the number of Non-Moving violation citations issued per month, again with the exception of 0-20% Minority tracts, which experienced about half the number of Non-Moving violations compared to other tract categories over this time period.

194. In light of the data reflected in Figures 3b and 3c, I conclude that the vast majority of the citation disparities (as reflected in Figure 3a) between higher Minority tracts and lower

Minority tracts that arose after the creation of the Strikeforce were with respect to Non-Moving violations.

195. In Figures 4a – 4f, I look at these Non-Moving citations in more detail. These figures present the data for mean monthly citation counts by tract type across the five-eras for the six most frequent types of Non-Moving violations: excessively tinted windows, other outwardly observable equipment violations (e.g., broken light, cracked windshield, worn tires, obstructed view, hard-to-read license plate), improper use of restraints, license/registration infractions, insurance infractions, and inspection infractions. Looking first at the average number of monthly tinted windows per tract in Figure 4a, we can see that during the Pre-Strikeforce era, there were exceedingly few citations for tinted windows in any type of tracts (an average of less than two such citations per month per tract during this era), with no significant disparity in such citations across tract types. Moving to the Strikeforce and BTVA eras, we can see that tinted window citations in 0-20% Minority tracts stayed largely constant and increased to an average of nearly 4 per month in the 20-40% Minority tracts. But tinted windows citations increased in the higher Minority tracts to a far greater extent. As can be seen, in the 60-80% Minority tracts, the average number of tinted windows citations went from 1.4 per month Pre-Strikeforce to almost 11 per month. In the 80-100% Minority tracts, they went from an average of 1.2 per month Pre-Strikeforce to an average of 16 per month—a more than thirteen-fold increase. Citations for tinted windows dropped substantially in all census tract categories in the Post-Strikeforce era. Even then, however, there were still an average of two to five times as many such citations per month in the 21 tracts that were 80-100% Minority, relative to the 19 tracts that were 0-20% Minority and the 21 tracts that were 20-40% Minority tracts, respectively. The last two sets of bars in Figure 4a show that tinted window citations again became infrequent in all tract categories during the Current era.



196. Figures 4b and 4c show that while smaller in number (see change in vertical axis relative to Tinted Windows), similar patterns arise with respect to "Other Observable Equipment" violations (Broken Light, Obstructed View, Worn Tires) and Seatbelt/Childseat citations as were observed with respect to Tinted Window citations in Figure 4a. Namely, there were a small and roughly equal number of such citations across all tract categories in the Pre-Strikeforce era. During the Strikeforce era, however, the number of such citations began to vary dramatically depending on the racial composition of the tract, with far more citations of these types being issued in the 60-80% Minority and 80-100% Minority tracts than in the 0-20% Minority and 20-40% Minority tracts. Those disparities persisted through the BTVA era, and to some extent into the Post-Strikeforce era before subsiding in the Current era.

Figure 4b: Mean Monthly Other Observable Equipment Citations per Tract

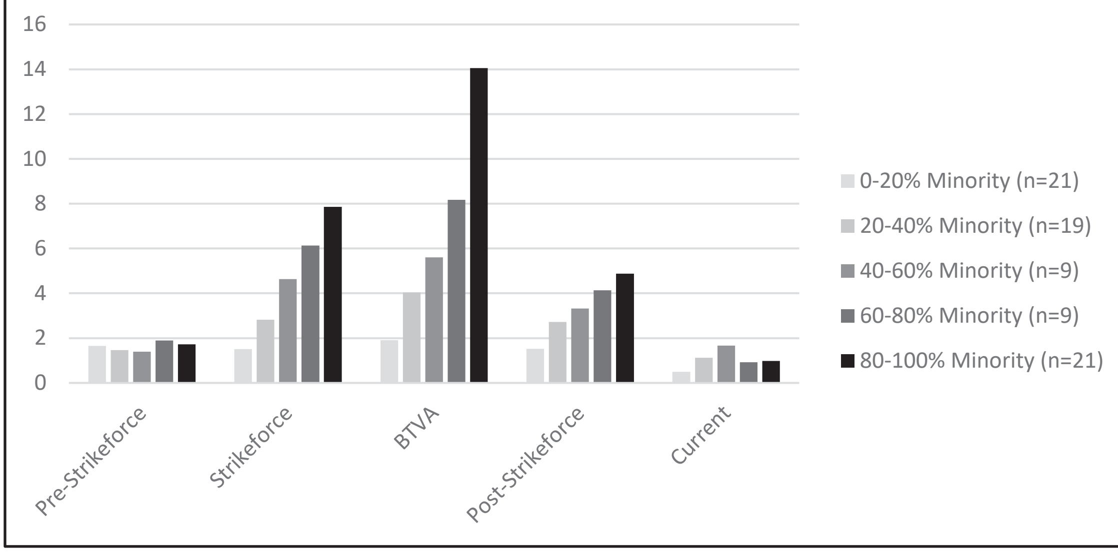
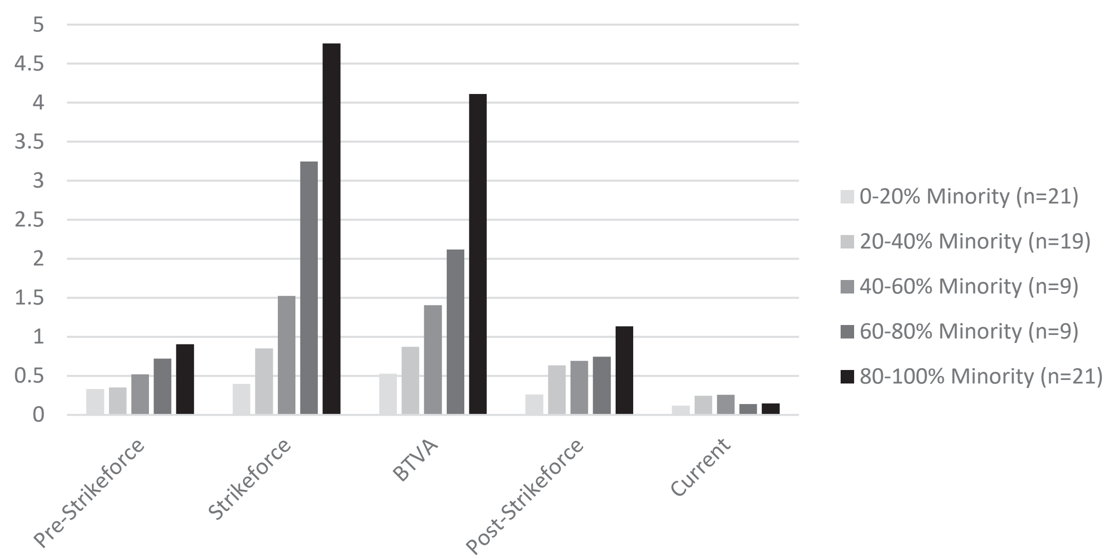
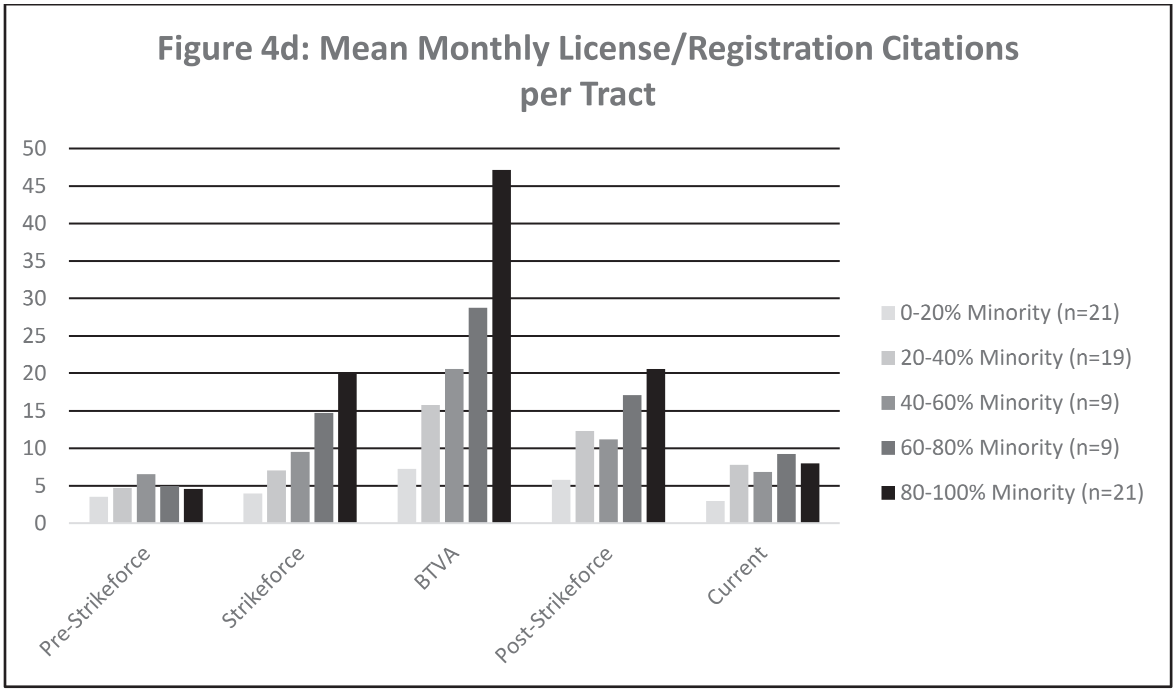


Figure 4c: Mean Monthly Restraint Citations per Tract



197. Figure 4d shows that during the Pre-Strikeforce era, there were also very similar numbers of License/Registration citations issued per month per tract across all tract categories, with an average of four to five such citations being issued per tract per month in each of the tract categories. As with the citation categories discussed in the previous figures though, starting in the Strikeforce era, License/Registration citations abruptly diverged between tracts with low Minority populations and those with high Minority populations. Two to three times more such citations were

issued per month in the 60-80% Minority and 80-100% Minority tracts relative to the 20-40 % Minority tracts, and four to five times more than in the 0-20% Minority tracts. These discrepancies in License/Registration citations grew even starker during the BTVA era, in which 60-80% Minority tracts experienced an average of almost 30 such citations per month and the 80-100% Minority tracts experienced an average of almost 50 such citations per month—a ten-fold increase from the Pre-Strikeforce era.

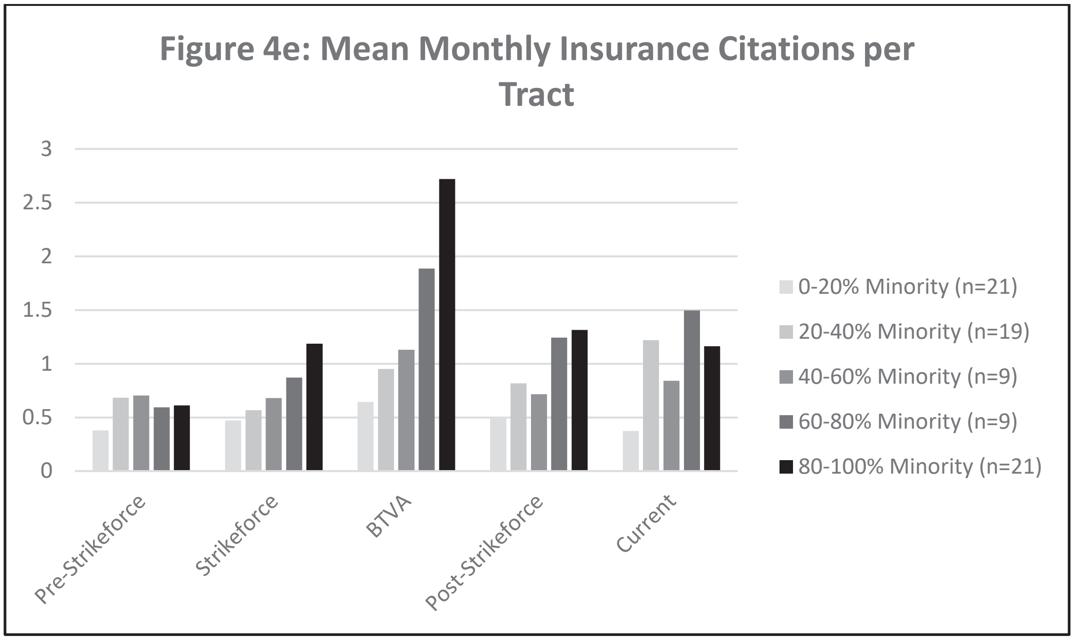


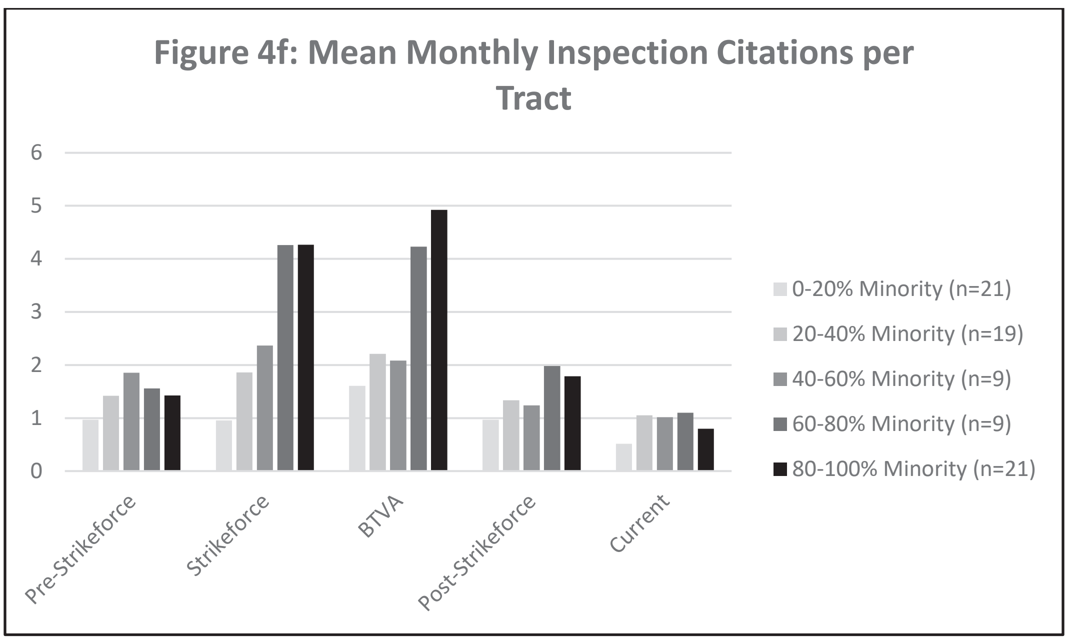
198. License/Registration citation counts by tract type are similar between the Strikeforce and Post-Strikeforce eras, with well over twice as many License/Registration citations issued per month in the higher Minority tracts as in lower Minority tracts during both eras.

199. During the Current era, the magnitude of discrepancies between the higher Minority tracts and the lower Minority tracts in License/Registration citations issued per month declined in comparison to the previous three eras. These disparities were not eliminated, however, as such citations became extremely rare only in the 0-20% Minority tracts.

200. Figures 4e and 4f show that similar patterns arose with respect to Insurance and Inspection citations as with respect to License/Registration citations, albeit at much lower levels. Namely, there were very similar numbers of such citations across all tract categories in the Pre-Strikeforce era and substantially greater numbers of such citations in the higher Minority tracts relative to the lower Minority tracts during the Strikeforce, BTVA, and Post-Strikeforce eras, with

such discrepancies waning during the Current era. As these figures illustrate, however, even in the Current era, far fewer citations of these types were issued in 0-20% Minority tracts than in tracts with greater Minority populations.





201. The above analysis reveals that although citation issuances were similar across lower and higher Minority tracts in the Pre-Strikeforce era, there was a large increase in citations for several sub-categories of Non-Moving infractions following the creation of the Strikeforce in July 2012, with almost all of the increase in such citations occurring in tracts with 60-100%

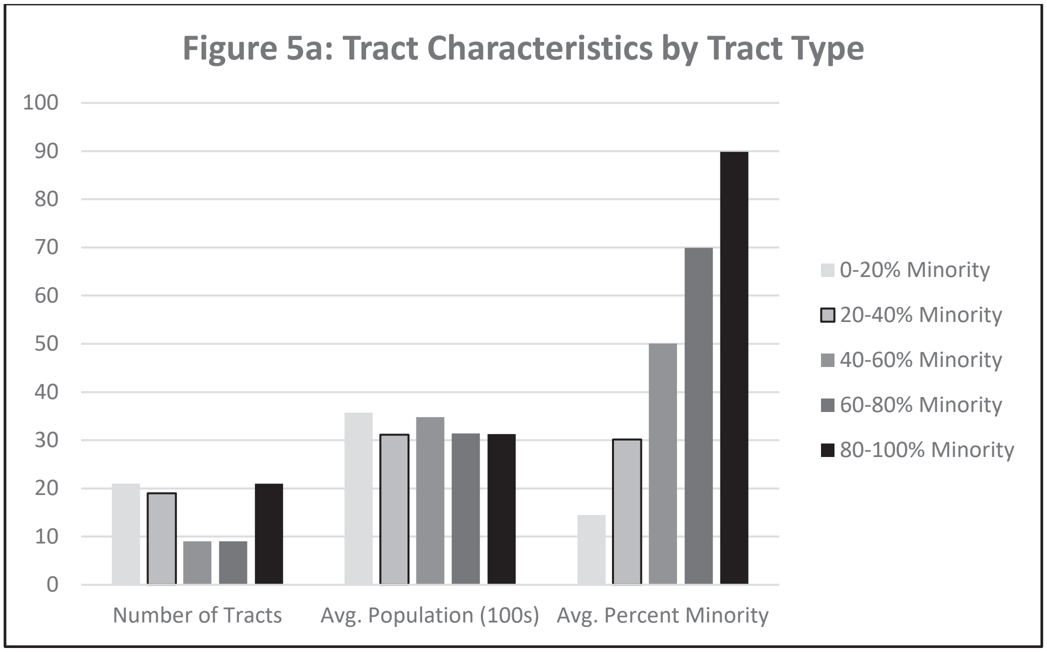
Black/Hispanic residents. As stated previously, I do not have data on the total number of actual equipment infractions that occurred by tract over time and do not think such data exists. However, it is implausible to me that the increase in citation issuance in High-Minority tracts during this period results from a change in the underlying behavior—*i.e.*, that there was an enormous increase in the rate at which drivers drove with equipment violations or without licenses, registration, or insurance, in High-Minority tracts in the Strikeforce, BTVA, and Post-Strikeforce eras relative to the Pre-Strikeforce era. Rather, this divergence in citations for equipment violations following June 2012 more plausibly resulted from a change in policing policies and/or practices around this time that focused efforts on ticketing such infractions in higher Minority neighborhoods.

202. The above descriptive analysis reveals that there was a large divergence in monthly citation counts between higher Minority tracts and lower Minority tracts starting in the Strikeforce era and continuing through the Post-Strikeforce era. As discussed in the methods section above, it is conceivable that the divergence in citations following the creation of the Strikeforce might have arisen not because the BPD targeted neighborhoods that have a larger Minority population, but because they targeted areas that experienced more accidents or crime events, which often correlated with areas in higher Minority neighborhoods. Greater police presence in these areas thus might account for higher number of citations in higher Minority neighborhoods. In the subsections that follow, I analyze this issue using several different methods, each of which supports the conclusion that accidents and crime cannot fully explain the racial disparities in citations.

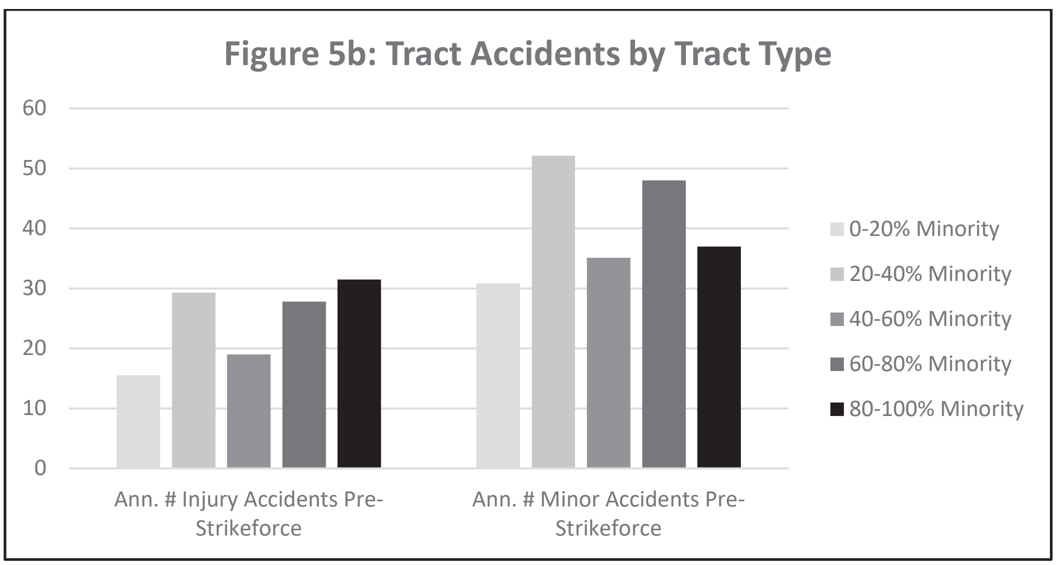
b) Racial Disparities in Citations Across Neighborhoods by Time Period (Controlling for Accidents and Crime)

203. As a first step in considering whether varying crime and accident rates might explain the differences in citations in different neighborhoods, I analyzed the crime and accident rates by tract. Figures 5a – 5c show how a variety of characteristics differ (or not) between tract type categories.

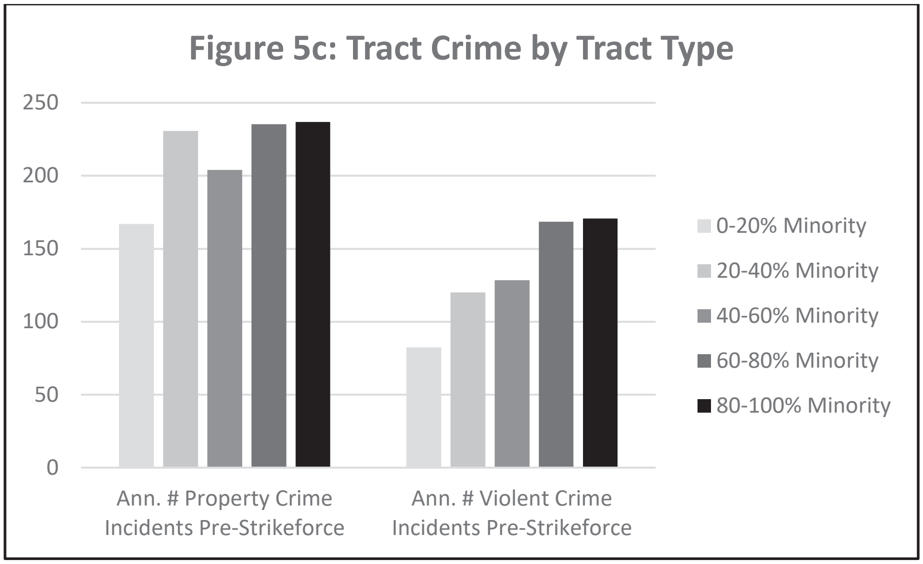
204. As alluded to earlier, Figure 5a shows that the average population across tract categories is quite similar, but the racial composition of tracts can differ dramatically.



205. Figure 5b shows that prior to the creation of the Strikeforce, the 80-100% Minority tracts had the most injury accidents, and the 20-40% Minority tracts had the second-most injury accidents. There appears to be no strong correlation between minor accidents and the percentage of Minorities in a tract.



206. Figure 5c shows that higher Minority tracts had higher numbers of property crime incidents than the lowest Minority tracts, and that 60-80% Minority and 80-100% Minority tracts had notably more violent crime incidents in the year leading up to the creation of the Strikeforce than both 0-20% Minority tracts and 20-40% Minority tracts. 40-60% Minority tracts were in the middle with respect to both property crime incidents and violent crime incidents.



207. Given the findings in Figure 5c, it is important to examine how citation counts differ across tract types after controlling for such differences. I do so in the following sub-sections.

208. Before moving on to the actual analyses, however, as discussed in the Methodology section above, for the following analyses I group census tracts into just three groups based on their racial make-up: High-Minority tracts (60-100% Minority), Low-Minority tracts (0-40% Minority), and Mixed-Race tracts (40-60% Minority). As discussed above, I use this categorization method for the following reasons. First, given several of my methods rely on comparing members of just two groups, categorizing tracts in the manner described above is useful as 70 of the 79 tracts lie in either High-Minority or Low-Minority categories. Therefore, these two categories have meaningful sample sizes within them for my analyses. Second, High-Minority and Low-Minority tracts as defined above are highly stable over time in terms of the racial make-up of their component tracts. Only one tract would fall into a different category at a different point in time if one were to use yearly ACS data. This tract is one of the 9 tracts categorized as Mixed-Race in my classification system.

c) Analysis of High Violent Crime Tracts

209. This section analyzes monthly citation counts in "High Violent Crime" tracts ("HVC tracts"), which I define to be in the top third of tracts by violent crime incidents in the 12 months prior to the implementation of the Strikeforce. This analysis shows that in the Strikeforce and BVTA eras, High-Minority HVC tracts received more than twice as many citations as Low-Minority HVC tracts. This provides evidence against the hypothesis that neighborhood crime

explains the much higher citation activity in higher Minority Tracts relative to lower Minority tracts.

210. Table 3 compares the distribution of Pre-Strikeforce violent crime in the High-Minority tracts (60-100% Minority) to the Low-Minority tracts (0-40% Minority) in this HVC group. As can be seen, there are over three times as many High-Minority tracts compared to Low-Minority tracts in this HVC group. However, violent crime incidents in the year leading up to the implementation of the Strikeforce are quite similar between High-Minority and Low-Minority tracts in this HVC group.

Table 3: Distribution of Violent Crime in High-Crime Tracts

	High-Minority (60-100%)	Low-Minority (0-40%)
Number of Tracts	18	5
Violent Crime Incidents (12-mo. Prior to Strikeforce)		
Mean	208.2	258.2
Median	186	188
25th Percentile	163	165
75th Percentile	241	242

High-Crime tracts include only census tracts in top third of violent crime incident distribution for the 12 months prior to Strikeforce.

211. Table 4 shows mean monthly citation counts in High-Minority HVC tracts and Low-Minority HVC tracts in the five different eras that I have discussed above. The top panel of Table 4 presents data for all types of citations, while the second and third panels present data for Moving violations and Non-Moving violations. The two left-most columns of numbers respectively show mean monthly citation counts in the High-Minority and Low-Minority tracts in this HVC subsample of tracts in each era. The right three columns show (i) the difference in mean monthly citation counts between High-Minority and Low-Minority tracts, (ii) the associated t-stat associated with the null hypothesis that the difference is zero; and (iii) the p-value associated with

the null hypothesis that the difference is zero.[27] Standard errors clustered by census tract are shown in parentheses and are used in calculating t-stat.

Table 4: Mean Monthly Citations per Tract (Pre-Strikeforce Highest Violent Crime Tracts Only)

	Tract Type				
Era	High-Minority	Low-Minority	Difference	t-stat	p-val
All Citations					
Pre-Strikeforce	17.9	26.9	**-9.0**	-1.91	**0.07**
	(2.7)	(3.87)	(4.72)		
Strikeforce	77.8	39.2	**38.5**	2.31	**0.03**
	(15.62)	(5.84)	(16.68)		
BTVA	126.7	60.8	**65.9**	2.29	**0.03**
	(25.96)	(12.48)	(28.8)		
Post-Strikeforce	54.1	50.7	**3.3**	0.18	**0.86**
	(8.23)	(16.73)	(18.65)		
Current	19.5	34.5	**-15.0**	-1.04	**0.31**
	(2.8)	(14.21)	(14.48)		
Moving Violations					
Pre-Strikeforce	5.2	9.2	**-4.0**	-1.67	**0.11**
	(.91)	(2.19)	(2.37)		
Strikeforce	12.3	10.3	**2.0**	0.58	**0.57**
	(2.55)	(2.27)	(3.42)		
BTVA	19.0	13.0	**6.0**	1.22	**0.23**
	(3.32)	(3.6)	(4.9)		
Post-Strikeforce	14.1	17.0	**-2.9**	-0.39	**0.70**
	(1.68)	(7.36)	(7.55)		
Current	6.5	11.0	**-4.4**	-0.91	**0.37**
	(.82)	(4.8)	(4.87)		
Non-moving Violations					
Pre-Strikeforce	12.7	17.7	**-5.0**	-1.58	**0.13**
	(1.92)	(2.56)	(3.2)		
Strikeforce	65.6	28.9	**36.7**	2.66	**0.01**
	(13.19)	(3.98)	(13.78)		
BTVA	107.8	47.8	**60.0**	2.42	**0.02**
	(22.7)	(9.98)	(24.8)		
Post-Strikeforce	39.8	32.6	**7.2**	0.62	**0.54**
	(6.68)	(9.51)	(11.62)		
Current	12.9	22.4	**-9.6**	-1.00	**0.33**
	(2.16)	(9.31)	(9.56)		

Includes only census tracts in top third of violent crime incident distribution for the 12 months prior to Strikeforce. Standard errors in parentheses, all clustered by census tract to take account of non-independence.

[27] A t-stat (or t-statistic) is the ratio of the difference in two sample means relative to the implied standard error of this difference. The t-stat can then be evaluated with respect to Student's t-distribution to determine the p-value, which is the likelihood that the observed difference could arise even if the two sample means are drawn from the same underlying distribution.

212. As can be seen in Table 4, in the Pre-Strikeforce era, there were on average about 9 fewer citations per month issued in the High-Minority HVC tracts relative to the Low-Minority HVC tracts. While this difference is not statistically significant at the 5% level, it is statistically significant at the 10% level.[28]

213. By contrast, on average there were almost 40 more citations per month (almost twice as many) issued in the High-Minority HVC tracts relative to the Low-Minority HVC tracts in the Strikeforce era, and over 65 more citations per month (more than twice as many) in the High-Minority HVC tracts relative to the Low-Minority HVC tracts during the BTVA era. Each of these differences is statistically significant at the 5% level.

214. In the Post-Strikeforce and Current eras, there were not statistically significant differences in the average number of citations issued per month in High-Minority HVC tracts relative to Low-Minority HVC tracts.

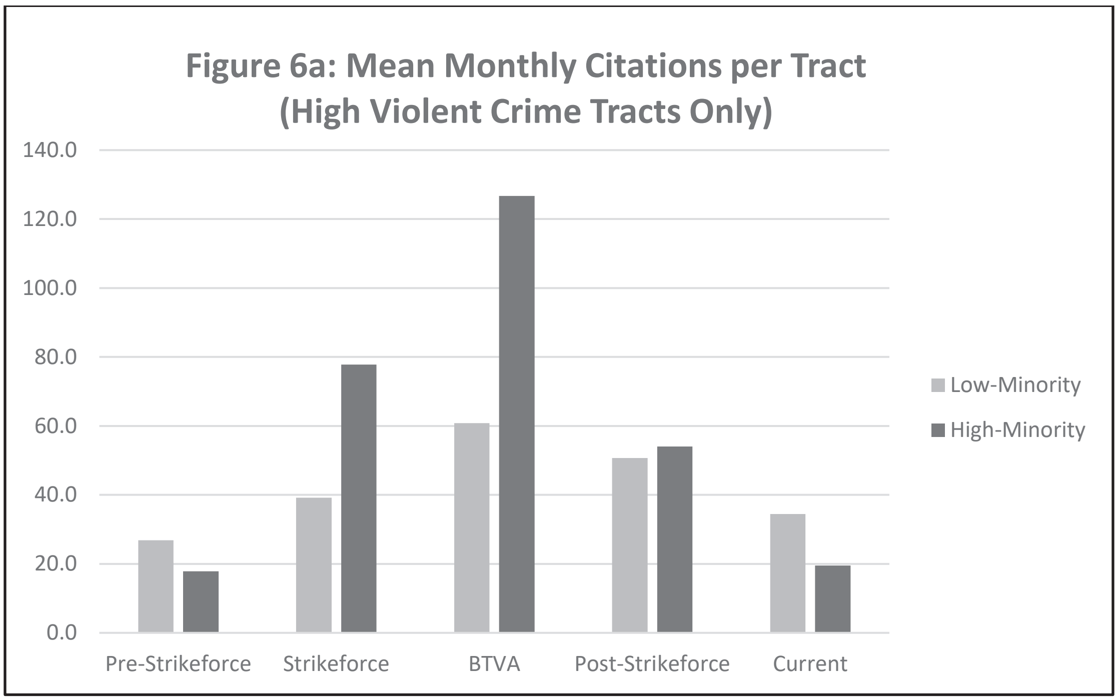
215. The middle panel of Table 4 presents the results of a similar exercise as that above, but looking only at citations for Moving violation citations. As can be seen, in none of the eras examined here was there a statistically significant difference in the average number of Moving violation citations issued per month between High-Minority HVC tracts and Low-Minority HVC tracts.

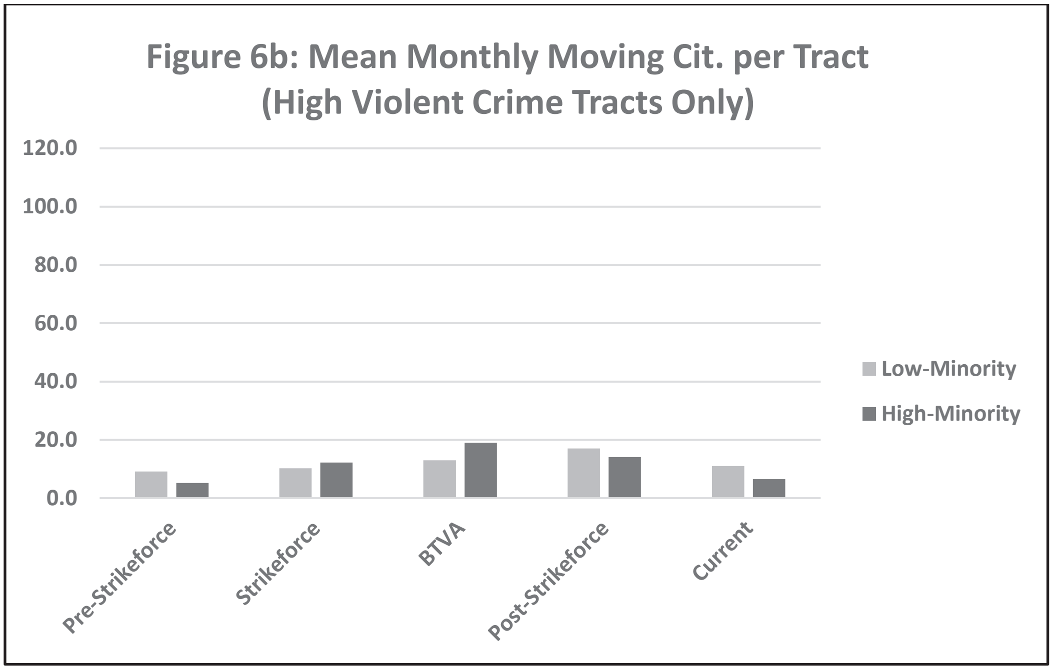
216. The bottom panel of Table 4 presents similar data for citations for Non-Moving violation citations. It shows that in the Pre-Strikeforce era, there were also not statistically significant differences in Non-Moving violation citations between the High-Minority HVC tracts and the Low-Minority HVC tracts. By contrast, during the Strikeforce era there were on average over 35 more Non-Moving violation citations issued per month in High-Minority HVC tracts than were issued per month in Low-Minority HVC tracts. Likewise, during the BTVA era there were on average almost 60 more Non-Moving violation citations issued per month issued in High-Minority HVC tracts than in Low-Minority HVC tracts. To summarize, in both the Strikeforce and BTVA eras, the BPD issued Non-Moving violation citations in High-Minority HVC tracts at more than twice the rate it issued such citations in Low-Minority tracts. Both of these differences are

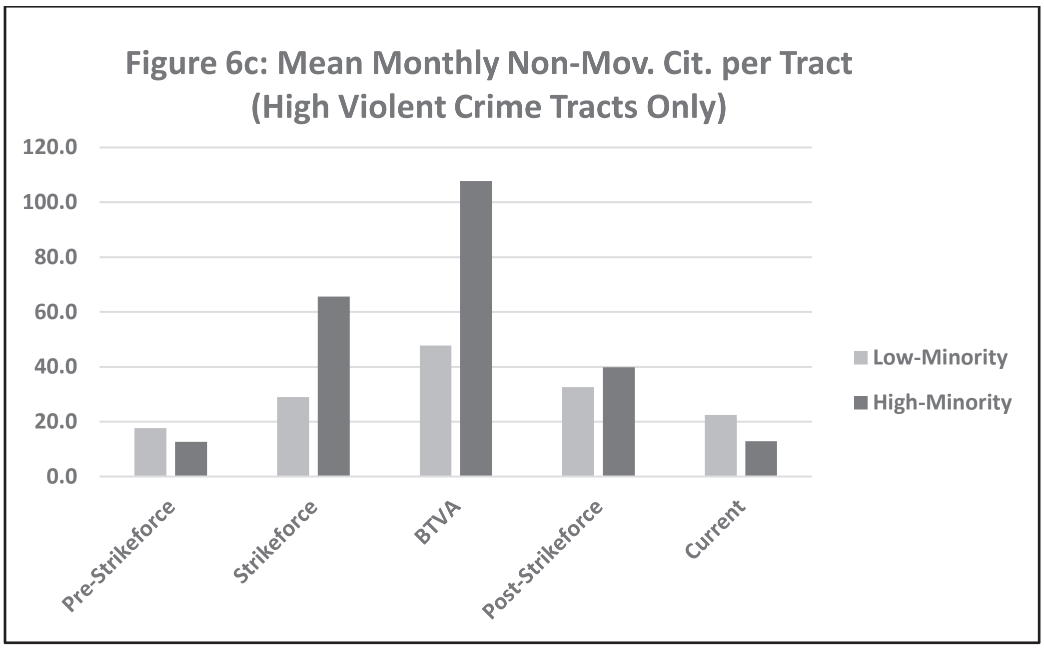
[28] Statistical significance at the 5% level is equivalent to having a p-value of 0.05 or less (see above footnote for discussion of how to interpret p-values). Statistical significance at the 10% or 1% levels is analogous.

statistically significant at the 5% level. In the Post-Strikeforce and Current eras, there were again no statistically different numbers of citations for Non-Moving violations per month between High-Minority HVC tracts and Low-Minority HVC tracts.

217. Figures 6a–6c summarize the results shown in Table 4 in graphical form. Overall, this exercise shows that when I restrict our analysis to only those tracts with the highest rates of violent crime in the 12 months leading up to the creation of the Strikeforce, there are very large discrepancies in monthly citations counts in the Strikeforce and BTVA eras, with High-Minority HVC tracts receiving on average well over double the number of citations per month compared to the Low-Minority HVC tracts. These discrepancies arise due to the increased rate at which BPD issued citations for Non-Moving violations. This analysis supports my conclusion that crime does not fully explain the racial discrepancies in citations across neighborhoods.







d) Multi-Variate Linear Regression Prediction Results

218. This section presents an analysis using the linear regression prediction techniques discussed above to examine how citation patterns differed between High-Minority and Low-Minority neighborhoods in Buffalo between 2012 and 2022 after controlling for accidents and crime in each neighborhood. My results again reveal that higher Minority neighborhoods experienced substantially and statistically significantly more citations, primarily for Non-Moving

violations, following the implementation of the Strikeforce until the onset of the COVID-19 pandemic, and that these disparities cannot be accounted for by differential rates of accidents or criminal incidents in higher Minority neighborhoods relative to lower Minority neighborhoods.

219. As discussed in the methods section above, for each era, in this analysis I employ a two-step process. In the first step I use Ordinary Least Squares (OLS) to regress monthly citations in each Low-Minority tract on the number of accidents (minor and injury) in each tract in the 12 months prior to the creation of the start of the relevant era, and the number of crime incidents (violent and property) in each tract in the 12 months prior to the start of the era.[29] In the second step, I use the coefficients obtained from these regressions to predict the number of citations for each High-Minority (and Mixed-Race) tract based on their number of accidents and crime incidents that occurred in each of these tracts in the 12 months prior to the start of the era. This gives me a measure of expected monthly citations for each of these tracts under the hypothesis that relationship between accidents and crime and subsequent citations is the same in High-Minority and Mixed-Race tracts as in Low-Minority tracts. I then take the difference between actual citations and predicted citations for each tract and see how large such differences are on average for High-Minority and Mixed-Race tracts.

220. The results of my regression analysis are shown in Tables 5a (all citation types), 5b (Non-Moving citations), and 5c (Moving citations).

221. The top panel of Table 5a shows the coefficients obtained from the prediction regressions using only the sample of Low-Minority tracts in each era. As can be seen, at least for the first four eras, among the five variables I include in the regressions, the most consistent predictor of citations in Low-Minority tracts is the number of violent crime incidents in the 12 months prior to the start of the era. Overall, the R-squared statistic for each of these regressions

---

[29] Note that I use only violent and property crime incidents as controls. I chose not to include drug crime and quality of life crimes as further controls. I used this approach because violent and property crime incidents are generally reported by a distinct victim. Drug crimes and quality of life crimes do not necessarily have identifiable victims who report these incidents to police, rather such offenses are more likely to become incidents due to police intervention. This means the number of these crime incidents in a given area is likely endogenous to police presence and activity in a given area. In other words, greater police presence in a neighborhood is likely to increase the rate at which drug and quality-of-life offenses are detected, to a much greater extent than the rate (if any) at which increased police presence causes greater reporting of violent and property crimes. As a result, these other categories of crime are not good controls for this analysis.

ranges from 0.29 to 0.44, suggesting that about one-third of the variation in citation counts across Low-Minority tracts is accounted for by the included tract characteristics.

Table 5a - Difference Between Actual and Expected Total Citations per Month (Pre-Era Controls)

	Coefficients on Predictors of Total Monthly Citations in Low-minority Tracts (OLS)				
Variable	Pre-Strikeforce Era	Strikeforce Era	BTVA Era	Post-Strikeforce Era	Current Era
Total Population (100s)	0.09	0.07	-0.23	-0.53***	-0.18
	(0.10)	(0.11)	(0.14)	(0.20)	(0.11)
Pre-era minor accidents (10s)	-2.19**	-2.83**	-2.36	5.62***	0.22
	(0.90)	(1.31)	(1.71)	(2.03)	(0.83)
Pre-era injury accidents (10s)	4.33**	4.17***	4.69*	-6.30	2.33
	(1.62)	(1.35)	(2.51)	(4.46)	(2.16)
Pre-era violent incidents (100s)	7.59**	10.17***	33.50***	18.61**	10.89*
	(2.81)	(2.69)	(7.40)	(8.50)	(5.60)
Pre-Era property incidents (100s)	-0.90	2.23	-5.29	3.32	1.92
	(1.78)	(1.48)	(4.15)	(4.38)	(1.99)
Constant	4.71	1.84	11.54*	16.83**	3.88
	(3.69)	(3.66)	(5.78)	(7.14)	(4.48)
Observations	240	1,440	1,240	1,000	1,360
R-squared	0.27	0.33	0.39	0.47	0.39
	Avg. Difference Between Actual and Predicted Monthly Citations by Tract Type				
High-minority Tracts (N = 30)	-7.49***	29.38***	37.68**	10.93*	-6.87***
	(2.12)	(9.59)	(15.95)	(5.82)	(2.33)
Mixed-race Tracts (N = 9)	1.25	11.56*	10.80	1.06	3.99
	(5.42)	(6.91)	(11.94)	(6.01)	(4.12)
Intercept (Low-Minority Tracts)	0.00	-0.00	0.00	-0.00	0.00
	(1.24)	(1.19)	(1.78)	(2.11)	(1.34)

Heteroskedastic robust standard errors clustered by census tract shown in parentheses. Accidents and criminal incidents correspond to the numbers of such incidents in the year prior to the start of each era. ***indicates significance at 1% level, **indicates significance at 5% level, *indicates significance at 10% level.

222. The bottom panel of Table 5a shows the key results for this analysis, capturing the difference between actual and predicted monthly citations per tract by tract type. As the first pair of rows in this bottom panel shows, among High-Minority tracts in the Pre-Strikeforce era, there were slightly but significantly *fewer* monthly citations than would have been predicted given recent crime and accident rates. However, as shown in the second and third columns, during the Strikeforce and BTVA eras, actual citations exceeded expected citations in High-Minority tracts by an average of 30–40 monthly citations per tract, with these differences being statistically significant at 1% level for the Strikeforce era and the 5% level for the BTVA era.

223. As can be seen in the fourth column of numbers in the lower panel of Table 5a, results are smaller for the Post-Strikeforce era, with the difference between actual and predicted monthly citations in High-Minority tracts being about 10 and this difference being statistically

significant at the 10% level. In the Current era, this difference between actual and predicted monthly citations is again negative (-6.87) and statistically significant at the 1% level.

224. The second pair of rows of the lower panel of Table 5a shows the results for Mixed-Race tracts. In general, the results are more muted than they are with respect to High-Minority tracts, but caution should be taken in putting too much confidence in these results due to the large standard errors resulting from the small number of observations (n=9).

225. The last pair of rows of the lower panel of Table 5a confirm that on average, there is no difference between actual and predicted citation levels in Low-Minority tracts, which should be true by construction in this statistical procedure.

226. Given the very different results between Moving Violations and Non-Moving violations discussed earlier, I also performed the above analysis separately for Non-Moving violation citations (Table 5b) and Moving violation citations (Table 5c).

227. Looking first at Non-Moving violation citations, the top panel of Table 5b shows that violent crime incidents in the 12 months prior to the start of the era are again the strongest predictor of Non-Moving violation citations in Low-Minority tracts.

Table 5b - Difference Between Actual and Predicted Monthly Non-moving Citations					
	Coeff. on Predictors of Monthly Non-moving Citations in Low-minority Tracts (OLS)				
Variable	Pre-Strikeforce Era	Strikeforce Era	BTVA Era	Post-Strikeforce Era	Current Era
Total Population (100s)	-0.01	0.02	-0.13	-0.30***	-0.12*
	(0.06)	(0.10)	(0.11)	(0.09)	(0.06)
Pre-era minor accidents (10s)	-1.67**	-2.54**	-2.27	2.59	-0.26
	(0.69)	(1.12)	(1.41)	(1.67)	(0.50)
Pre-era injury accidents (10s)	2.37**	3.04**	3.78*	-3.16	2.33*
	(1.11)	(1.13)	(2.03)	(3.06)	(1.32)
Pre-era violent incidents (100s)	8.57***	9.61***	28.26***	15.80***	9.45**
	(2.03)	(2.20)	(7.02)	(5.00)	(3.68)
Pre-Era property incidents (100s)	-1.72	1.08	-4.43	0.72	0.96
	(1.50)	(1.09)	(3.81)	(1.55)	(1.08)
Constant	5.14*	2.25	5.64	8.08***	-0.00
	(2.57)	(3.37)	(4.31)	(2.92)	(2.22)
Observations	240	1,440	1,240	1,000	1,360
R-squared	0.31	0.29	0.39	0.50	0.44
	Avg. Diff. Between Actual and Predicted Monthly Non-moving Citations				
High-minority Tracts (N = 30)	-4.74***	27.12***	35.01**	10.41**	-4.76***
	(1.38)	(8.09)	(13.99)	(4.39)	(1.69)
Mixed-race Tracts (N = 9)	-0.44	9.25*	10.82	3.35	3.11
	(3.04)	(5.18)	(10.61)	(3.73)	(2.14)
Intercept (Low-Minority Tracts)	-0.00	0.00	0.00	0.00	-0.00
	(0.84)	(0.99)	(1.49)	(1.08)	(0.69)

Heteroskedastic robust standard errors clustered by census tract shown in parentheses. Accidents and criminal incidents correspond to the numbers of such incidents in the year prior to the start of each era. ***indicates significance at 1% level, **indicates significance at 5% level, *indicates significance at 10% level. Non-moving violation citations include citations for equipment infractions (e.g., tinted windows, obscured view, seatblelt/childseat infractions, worn tires, broken lights), license/registration infractions, insurance infractions, and inspection infractions.

228. The top row of the lower panel of Table 5b shows that in the Pre-Strikeforce era, on average, there were fewer actual Non-Moving violation citations per month than predicted in High-Minority tracts. However, in the Strikeforce and BTVA eras, there were on average 27 and 35 more Non-Moving violation citations per month in High-Minority tracts than would have been predicted, and both of these values are statistically significant at the 5% level or higher. This difference between actual and predicted Non-Moving violation citations in High-Minority tracts falls in the Post-Strikeforce era to around 10 per month, but this is still significantly different than zero at the 5% level. In the Current era, the average difference between actual and predicted Non-Moving violation citations in High-Minority tracts is again negative and statistically significant, though very small in magnitude. Results for Mixed-Race tracts are much smaller in magnitude and again not very precisely estimated due to small sample size.

229. Table 5c shows the results for Moving violation citations. As the top two rows of the bottom panel illustrates, these data suggest a similar general pattern as with respect to Non-Moving citations, with High-Minority tracts receiving fewer citations than would have been expected in the Pre-Strikeforce era, more citations from the Strikeforce era through the Post-Strikeforce era, and then similar or fewer citations in the Current era. The magnitudes of these disparities are considerably smaller, however, and are generally not statistically significant.

Table 5c - Difference Between Actual and Predicted Monthly Moving Citations					
	Coeff. on Predictors of Monthly Moving Citations in Low-minority Tracts (OLS)				
Variable	Pre-Strikeforce Era	Strikeforce Era	BTVA Era	Post-Strikeforce Era	Current Era
Total Population (100s)	0.10	0.05	-0.08	-0.10	-0.02
	(0.07)	(0.03)	(0.05)	(0.09)	(0.06)
Pre-era minor accidents (10s)	-0.52	-0.27	-0.44	1.91***	0.46
	(0.39)	(0.28)	(0.35)	(0.53)	(0.36)
Pre-era injury accidents (10s)	1.97**	1.12***	1.18*	-1.00	-0.09
	(0.85)	(0.36)	(0.59)	(1.37)	(0.96)
Pre-era violent incidents (100s)	-1.05	0.47	5.19***	0.79	1.15
	(1.50)	(0.90)	(1.54)	(3.77)	(2.15)
Pre-Era property incidents (100s)	0.87	1.21**	-0.17	2.85	1.24
	(0.76)	(0.52)	(0.73)	(3.71)	(0.93)
Constant	-0.52	-0.56	4.19**	2.98	1.69
	(2.38)	(0.93)	(1.84)	(3.19)	(2.51)
Observations	240	1,440	1,240	1,000	1,360
R-squared	0.16	0.27	0.17	0.28	0.18
	Avg. Diff. Between Actual and Predicted Monthly Moving Citations				
High-minority Tracts (N = 30)	-2.69**	2.42	2.60	1.53	-0.89
	(1.05)	(1.64)	(2.08)	(1.53)	(0.95)
Mixed-race Tracts (N = 9)	1.82	2.48	0.24	-0.18	0.93
	(2.90)	(1.82)	(1.81)	(2.81)	(1.64)
Intercept (Low-Minority Tracts)	0.00	0.00	0.00	0.00	0.00
	(0.66)	(0.35)	(0.48)	(1.04)	(0.71)

incidents correspond to the numbers of such incidents in the year prior to the start of each era. ***indicates significance at 1% level, **indicates significance at 5% level, *indicates significance at 10% level. Moving violation citations include moving violations (e.g., speeding, reckless driving, failure to stop at red light or stop sign), DUIs, and mobile electronics infractions.

230. As discussed in the Methodology section, I also conducted a version of this analysis using accidents and criminal incidents in the 12 months prior to the start of the Strikeforce, rather than the 12 months prior to the start of the relevant era, as the key conditioning variables. The results of this analysis are shown in Appendix Tables A1a-A1c and the results are very similar to those discussed above. Overall, these results show that from the implementation of the Strikeforce

through the start of the COVID-19 pandemic, the actual number of citations, particularly citations for Non-Moving violations, in High-Minority tracts significantly exceeded the number that would be predicted to have been issued in those tracts based on (i) the population, the number of accidents, and then number of criminal incidents in such tracts, and (ii) the relationship in Low-Minority tracts between those factors and the number of citations issued. In other words, these results again indicate that the higher monthly citation counts in High-Minority neighborhoods following the implementation of the Strikeforce through the onset of the COVID-19 pandemic do not simply reflect the BPD targeting citations at areas with higher populations or number of accidents or greater criminal activity, but rather represent differential treatment of High-Minority tracts relative to Low-Minority tracts.

e) Propensity Score Matching Results

231. This section presents the results of a propensity matching analysis with respect to monthly citation counts by racial composition of tracts. Similar to the regression prediction analysis above, this methodology attempts to assess whether there exist racial disparities in citations after controlling for other factors that may be correlated with racial composition of tracts but may also have directly affected citation rates themselves.

232. As discussed in the Methodology section above, these propensity score matching estimates compare citation counts for High-Minority tracts to a "matched" sample of Low-Minority tracts that had similar numbers of accidents and crime incidents in the year leading up to the creation of the Strikeforce. The tract characteristics I match on (which are referred to as conditioning characteristics) are the number of minor accidents and injury accidents that occurred in the tract between July 2011 and June 2012, and the number of violent crime incidents and property crime incidents that occurred in the tract between July 2011 and June 2012. Recall that I use Pre-Strikeforce conditioning characteristics for two reasons. First, to the extent BPD citation practices after the creation of the Strikeforce affected things like crime and accidents, controlling for them contemporaneously would be problematic. Second, this matching analysis uses the same conditioning characteristics for each era so that the matches are the same each era. This will allow for direct comparisons of results across eras that are not contaminated by changes in the sample. Controlling based on crime and accident rates that changed during the period in question would potentially change the sample.

233. The first two columns of numbers in Table 6 show how High-Minority tracts and Low-Minority tracts differed in these conditioning characteristics overall. As alluded to previously, High-Minority tracts and Low-Minority tracts differ substantially on certain dimensions—namely, High-Minority tracts have notably higher numbers of injury accidents, as well as higher numbers of violent and property crime incidents, than Low-Minority tracts in the 12 months leading up to the implementation of the Strikeforce.

Table 6: Comparison of Tract Characteristics

	All Tracts		Tracts in Common Support	
	High Minority	Low Minority	High Minority	Matched Low Minority
Minor Accidents	40	41	37	36
Injury Accidents	30	22	25	25
Vio. Crime Incidents	170	100	134	130
Prop. Crime Incidents	235	197	200	188
Tracts	30	40	18	10

Note: Accidents and criminal incidents correspond to the numbers of such incidents in the year leading up to the implementation of the Strikeforce (July 2011-June 2012). High Minority refers to tracts that are 60-100 percent Black and/or Hispanic, Low Minority refers to tracts that are 0-40 percent Black and/or Hispanic.

234. As is standard in propensity score matching, in order to do the match, I first collapse this multi-dimensional vector of conditioning characteristics into a uni-dimensional value via estimating a Probit regression, where I limit the sample to High-Minority and Low-Minority tracts, and regress the High-Minority indicator for each tract on these conditioning characteristics for each tract, and use these results to calculate a "propensity" that each tract is High-Minority (i.e., the propensity score or pscore).

235. In nearest neighbor matching, each High-Minority tract is then matched to the Low-Minority tract with the closest propensity score. However, as discussed in the Methodology section, in order to match "like-to-like", I limit the sample to those tracts within the "common support," or in other words, I limit the sample to tracts with a propensity score that lies within the range of the maximum and minimum propensity scores across the Low-Minority tracts. Limiting the sample to include only tracts with propensity scores within the common support excludes 12 of the 30 High-Minority tracts (all of which had propensity scores that exceeded the highest propensity score among the Low-Minority tracts).

236. Given the fact that there are some large differences between High-Minority tracts and Low-Minority tracts overall with respect to the conditioning characteristics as shown in the first two columns of Table 6, it is not surprising that some of the High-Minority tracts are excluded by the common support restriction. However, as the two right-most columns in Table 6 show, once the sample is restricted to the common support and matches for the remaining High-Minority tracts are made, High-Minority tracts are indeed quite comparable to the matched Low-Minority tracts in terms of Pre-Strikeforce crime and accident incidents. Indeed, I cannot reject the null hypothesis of equality between the common support High-Minority tracts and the matched tracts from the Low-Minority sample regarding each of the tract characteristics shown here.[30]

237. By comparing the first column of numbers in Table 6 to the third column of numbers in Table 6, we can see that the average number of violent and property crime incidents are notably larger in the full sample of High-Minority tracts than in just the High-Minority tracts that lie in the common support that are used for the matching estimates. The High-Minority tracts that fell outside the common support restriction were generally the tracts with highest rates of violent crime. Therefore, these matching results tend to compare citation rates of the relatively lower crime High-Minority tracts to the higher crime Low-Minority tracts. This latter point can be seen by comparing the second column of numbers to the fourth column of numbers in Table 6, where we can see that the Low-Minority tracts used in the matching process have higher numbers of injury accidents and more violent crime than Low-Minority tracts overall.

[30] Note that I do not use tract population as a conditioning variable in calculating the Propensity score. In general, in Propensity Score matching, the fewer conditioning variables there are, the better the match among these variables. Given average population is similar between High-minority and Low-Minority tracts (i.e., 3,131 and 3,357 respectively), I chose not to include population as a conditioning variable to increase match quality on those conditioning variables that do differ on average between High- and Low-Minority tracts (i.e., accidents and criminal incidents). However, I have examined what happens to the results if I do include tract population as a conditioning variable. When I include tract population as a conditioning variable, 3 additional High-minority tracts lie outside the common support, and therefore are not included in the analysis. Moreover, the "match" on criminal incidents is worse in the sense that the difference between the average number of criminal incidents in High-minority tracts in the common support and in the matched Low-Minority tracts is greater than in Table 6 when tract population is not included. However, the basic results coming from the matching exercise are similar to those shown in Table 7. In other words, if I were to include tract population as another conditioning variable in calculating Propensity scores, my basic conclusions would be unaffected.

238. Given the sample is restricted to those within the common support, which excludes some of the highest crime tracts, this analysis is focusing on a different set of tracts that the earlier analysis that focused on High Violent Crime (HVC) tracts. In words, this analysis focuses on uncovering whether there were significant citation disparities between High-Minority and Low-Minority tracts that are quite similar with respect to minor and injury accidents, and property and violent crimes. This analysis necessarily excludes some of the tracts with the highest rates of violent crime, as those tracts do not fall within the common support.

239. Table 7 shows the Propensity Score Matching results comparing average number of monthly citations in High-Minority tracts to matched Low-Minority tracts, by era. The top panel of Table 7 shows the results for the Pre-Strikeforce era, the second panel shows the analogous results for the Strikeforce era, the third panel shows the results for the BTVA era, the fourth panel shows the results for the Post-Strikeforce era, and the bottom panel shows the results for the Current era. Each panel shows the results for all citations, Non-Moving violation citations, and Moving violation citations separately. Each panel shows the unconditional (“Unmatched”) results, as well as the results coming from the Propensity Score Matching procedure (“Matched”).

Table 7: Propensity Score Matching Estimates of Citations

		Avg Cit./tract/mo					
Outcome Variable	Sample	High Minority	Low Minority	Diff.	Clustered S.E.	t-stat	p-val
Pre-Strikeforce (Jan 2012-June 2012)							
All Citations	Unmatched	15.37	15.19	0.18	2.50	0.07	0.941
	Matched	11.75	22.05	-10.31	3.60	-2.87	0.008
Non-moving Citations	Unmatched	10.81	9.04	1.76	1.74	1.01	0.315
	Matched	8.25	14.88	-6.62	3.32	-1.99	0.056
Moving Citations	Unmatched	4.61	6.14	-1.53	1.02	-1.50	0.139
	Matched	3.54	7.14	-3.60	1.20	-3.00	0.006
Strikeforce (July 2012-June 2015)							
All Citations	Unmatched	59.94	18.82	41.13	10.45	3.94	0.000
	Matched	37.64	24.87	12.78	5.41	2.36	0.025
Non-moving Citations	Unmatched	50.33	13.18	37.15	8.82	4.21	0.000
	Matched	31.34	17.89	13.44	4.76	2.83	0.009
Moving Citations	Unmatched	9.70	5.61	4.08	1.75	2.34	0.022
	Matched	6.35	6.88	-0.53	0.95	-0.56	0.582
BTVA (July 2015-Jan 2018)							
All Citations	Unmatched	97.32	29.00	68.32	17.37	3.93	0.000
	Matched	66.92	33.76	33.17	12.64	2.62	0.014
Non-moving Citations	Unmatched	81.95	21.29	60.66	15.16	4.00	0.000
	Matched	55.27	25.56	29.71	11.12	2.67	0.012
Moving Citations	Unmatched	15.39	7.29	8.10	2.36	3.44	0.001
	Matched	11.63	8.00	3.63	1.97	1.84	0.076
Post-Strikeforce (Feb 2018-Feb 2020)							
All Citations	Unmatched	45.65	27.40	18.25	6.80	2.68	0.009
	Matched	35.29	22.59	12.70	4.84	2.62	0.014
Non-moving Citations	Unmatched	32.88	15.08	17.81	4.99	3.57	0.001
	Matched	24.61	13.49	11.12	3.56	3.13	0.004
Moving Citations	Unmatched	12.54	10.27	2.28	2.06	1.10	0.274
	Matched	10.46	8.04	2.41	1.77	1.37	0.183
Current (March 2020-December 2022)							
All Citations	Unmatched	19.27	15.69	3.58	3.23	1.11	0.272
	Matched	17.04	11.67	5.37	2.70	1.99	0.057
Non-moving Citations	Unmatched	12.77	8.49	4.27	2.27	1.88	0.064
	Matched	10.96	6.15	4.81	1.93	2.49	0.019
Moving Citations	Unmatched	6.38	6.20	0.18	1.17	0.15	0.88
	Matched	5.94	4.66	1.28	1.24	1.03	0.31

Note: Standard errors are calculated to be robust to heteroskedasticity and clustered by census tract. Non-moving violations include citations for equipment infractions (e.g., tinted windows, obscured view, seatblelt/childseat infractions, worn tires, broken lights), license/registration infractions, insurance infractions, and inspection infractions. Moving-violations include citations for speeding, reckless driving, failure to stop at red light or stop sign, DUIs, failure to dim headlights, and mobile electronics infractions. High Minority refers to tracts that are 60-100 percent Black and/or Hispanic, Low Minority refers to tracts that are 0-40 percent Black and/or Hispanic.

240. As can be seen in the top row of the top panel of Table 7, during the Pre-Strikeforce era, the mean monthly citations issued in High-Minority tracts relative to Low-Minority tracts were not statistically different in the Unmatched sample (the rows labeled "Unmatched"). However, in the Matched sample (the rows labeled "Matched"), mean monthly citations were actually slightly lower in the High-Minority tracts than the matched Low-Minority tracts. The lower parts of this top panel show that this same pattern is also true if we examine Non-Moving violation citations and Moving violation citations separately.

241. The second panel of Table 7 shows the results for the Strikeforce era. The top row in the second panel shows that when comparing all High-Minority tracts to all Low-Minority tracts (i.e., the "Unmatched" sample), average monthly citations during this era were well over three times as high in High-Minority tracts than in Low-Minority tracts (which is similar to what was shown in Figure 3a). Of particular interest for this section, however, is the second row in this panel, which shows the results for the "Matched" sample, where we only include High-Minority tracts within the common support and compare them to "matched" Low-Minority tracts. Using the Matched sample, drivers in each High-Minority tract on average received over 13 more citations per month than did drivers in each Low-Minority tract that had similar accident and crime levels in the year prior to the formation of the Strikeforce. This difference is statistically significant at the 5% level.

242. As this second panel of Table 7 further demonstrates, during this Strikeforce era, the difference in monthly citations between High-Minority and Low-Minority tracts reflects differential citations for Non-Moving violations only. There is a large and statistically significant (at the 5% level) difference in the rate at which drivers received Non-Moving citations in the High-Minority and Low-Minority matched tracts (31.8 per tract per month vs. 18.07 per tract per month, respectively). Moreover, this difference in average Non-Moving violation citations between High-Minority and matched Low-Minority tracts is statistically significant at the 5% level.

243. The third panel in Table 7 shows the results for the BTVA era. Differences in average monthly citations between High-Minority tracts in the common support and matched Low-Minority tracts are even larger during this era, with an average of over 30 more citations issued per month in High-Minority tracts relative to matched Low-Minority tracts, and this result is statistically significant at the 5% level. Again, as can be seen in the lower rows of this third panel,

these citation differences across tract types during this era largely reflect disparities in Non-Moving violation citations rather than Moving violation citations.

244. The fourth panel in Table 7 shows that there continued to be statistically higher monthly citation counts in High-Minority tracts relative to matched Low-Minority tracts in the Post-Strikeforce era. Again, as can be seen in the lower rows of this panel, this was also driven by Non-Moving violations rather than Moving violations.

245. The bottom panel of Table 7 shows the results for the Current era. As can be seen, while differences in average monthly citation rates between High-Minority and Low-Minority tracts were quite a bit smaller than in earlier eras, both overall and in the matched sample, such differences were still generally statistically significant at the 10% level for all citations and at the 1% level for Non-Moving citations. In other words, within tracts with similar accident and crime levels, BPD continued to issue more citations in general and for Non-Moving violations within tracts with higher Minority populations than in similar tracts with lower Minority populations.

246. In summary, the results of this Propensity Score Matching analysis reveal that after the implementation of the Strikeforce, citations in High-Minority tracts diverged markedly from Low-Minority census tracts that were similar in terms of accidents and crime. Almost all of this divergence occurred with respect to Non-Moving violation citations. Moreover, these significant differences in monthly citations between High-Minority tracts and Low-Minority census tracts persisted into the Current era. Given these matching estimates are comparing High-Minority tracts to Low-Minority tracts that had similar pre-Strikeforce accident and criminal occurrences, these discrepancies cannot be explained by the BPD simply increasing citations in neighborhoods that experienced a large number of accidents or had a high number of violent and/or property crime incidents prior to the creation of the Strikeforce.

247. These results support a conclusion that is very consistent with that from the results from looking just at High Violent Crime (HVC) tracts and the Regression Prediction results discussed earlier—*i.e.*, that BPD began issuing citations in High-Minority neighborhoods much more intensively after June 2012, particularly with respect to citations for Non-Moving violations, and that this intensive ticketing largely persisted until 2020, after which it moderated or ceased, depending on how the data is analyzed. Moreover, this greater ticketing in higher Minority neighborhoods does not appear to be fully explained by higher crime levels or greater numbers of accidents in these higher Minority neighborhoods.

2. Analysis of Racial Disparities in Citations Across Individuals Within Neighborhoods

248. The above analysis focused on assessing racial disparities in the ticketing patterns *across* neighborhoods (as captured by census tracts). In this section, I analyze racial disparities across individuals in ticketing *within* neighborhoods, where again I use census tracts as my definition of neighborhoods.

a) Data on Race of Driver

249. As discussed in the Methodology section above, data on race of the cited individual comes from two sources: TraCS and OD. Table 8 shows the coverage of observations with race data from both these sources separately, as well as combined, by era. As can be seen, the fraction of citations for which we know driver race in the TraCS data ranged from a low of 52% in the Pre-Strikeforce era, to a high of 67% in the Post-Strikeforce era. With respect to the OD dataset, the fraction of citations with race data was quite high for the first four eras, but dropped substantially in the Current era. Overall though, between the two datasets, race data is available for between 78–83% of citation observations across eras, and 82% of citations overall.

Table 8 - Fraction of Obs with Driver Race Recorded by Era

Era	Tracs	OD	Tracs or OD
Pre-Strikeforce	0.52	0.64	0.79
Strikeforce	0.57	0.69	0.82
BTVA	0.60	0.68	0.83
Post-Strikeforce	0.66	0.60	0.82
Current	0.67	0.52	0.79
Overall	0.61	0.65	0.82

250. As stated above, my analysis uses TraCS data for driver race when that data is available. Only when that data is unavailable do I use the OD measure. However, for over 160,000 citations, I have race data from both sources. This allows me to cross-validate the accuracy of my race measure. Over 96% of the time both measures are the same.

251. In terms of the racial composition of cited drivers, the top row of Table 9 shows that among all those individuals for whom race data is recorded (from TraCS or OD), 76% of all citations are recorded as being issued to Minorities, and only 24% are issued Non-Minorities. By comparison, according to the ACS, the City of Buffalo was generally a little under 50% Minority over the time period examined here.

Table 9 - Driver Race Across Sources			
Sample	**Minority**	**Non-Minority**	**Obs**
1 - Valid Race (Tracs and/or OD)	0.76	0.24	306,603
	Tracs Minority	**Tracs Non-minority**	
2 - Valid Tracs Race (w/ or w/out OD Race)	0.76	0.24	228,041
3 - Valid Tracs Race (w/ valid OD Race)	0.84	0.16	163,575
4 - Valid Tracs Race (w/out valid OD Race)	0.57	0.43	64,466
	OD Minority	**OD Non-minority**	
5 - Valid Tracs Race (w/ valid OD Race)	0.85	0.15	163,575
6 - Tracs Race Missing (w/ valid OD Race)	0.75	0.25	78,562
Obs with Tracs and OD Race Missing	-	-	67,783

252. The second row of Table 9 shows that for citations for which race data is recorded in TraCS (*i.e.*, regardless of whether there is also race data recorded in OD), 76% are classified as Minority and 24% are classified as Non-Minority, which is the same as in the full sample of all citations for whom race is available. The third and fourth rows of Table 9 look at the race composition for citations (as recorded by TraCS) among those with race recorded in both TraCS and OD (row 3) and among citations with race recorded in TraCS but not OD (row 4). As can be seen, among citations with race recorded in TracCS and OD, about 84% are Minority according to TraCS. In comparison, among those who have a recorded TraCS race but no OD race, only 57% are Minority. This suggests that those citations with no OD race are less likely to be Minorities than those with OD race, though Minority drivers still make up a notable majority of citations even among those missing OD race/ethnicity.

253. The fifth and sixth rows of Table 9 look at the race composition as measured by OD, among those citations with recorded TraCS race and OD race (row 5) and among those citations with unrecorded TraCS race but with an OD race (row 6). Row 5 shows that among those with recorded TraCS race and OD race, about 85% are Minority according to OD. This is expected, as this is the same population as in row 4, but simply uses the OD race measure rather than the TraCS race measure, and these two records are generally the same. By comparison, among those who do not have a recorded TraCS race but do have an OD race, about 75% are Minority according to OD. This is similar to the overall fraction Minority among those with TraCS and/or OD race.

254. In summary, while we do not know for certain the Minority status of those for whom race information is missing from both data sources, this cross-validation exercise indicates that a substantial majority of those missing race data in just one of the sources are Minorities. However, as I discussed extensively in the Methodology section, there is a reasonable way to impute the racial composition of tickets issued to those of unrecorded race using data on racial make-up of zip code of cited individual's residence; the racial make-up of tract of issuance; the type of citation; the era of issuance; and whether the citation was issued to an individual from Buffalo, a first-ring suburb, or from another area.

255. Table 10 shows the imputed racial distribution of those for whom race is unrecorded using the procedure I laid out in the Methodology section. The first row of numbers shows the actual fraction Minority for those for whom race is recorded for All citations, Moving violations, and Non-Moving violations. As mentioned above, among those with recorded race, 76.1% of all citations with race recorded were issued to Minorities. However, as shown in Table 10, among those with recorded race, only 64.7% of citations for Moving violations were issued to Minorities, but 80.1% of Non-Moving violations were issued to Minorities.

Table 10 - Imputed Race Distribution for Drivers with Unrecorded Race			
	Percent Minority		
	All Citations	Moving Violations	Non-moving Violations
Citywide			
Race Recorded (Actual)	76.1%	64.7%	80.1%
Race Recorded (Imputed)	76.5%	64.6%	80.4%
Race Unrecorded (Imputed)	64.8%	52.2%	70.4%
High-Minority Tracts			
Race Recorded (Actual)	87.2%	80.7%	88.9%
Race Recorded (Imputed)	87.0%	80.1%	88.9%
Race Unrecorded (Imputed)	79.7%	70.2%	83.0%
Mixed-race Tracts			
Race Recorded (Actual)	65.0%	57.1%	67.9%
Race Recorded (Imputed)	65.6%	56.8%	68.6%
Race Unrecorded (Imputed)	59.8%	51.8%	63.2%
Low-Minority Tracts			
Race Recorded (Actual)	57.0%	46.7%	63.0%
Race Recorded (Imputed)	58.4%	47.2%	63.8%
Race Unrecorded (Imputed)	49.9%	39.9%	54.1%
*This table only reflects citations were could be mapped to a census tract within the city of Buffalo.			

256. The second row of Table 10 shows what my procedure would impute for the racial distribution for these drivers *with* recorded race. As can be seen, among cited drivers with recorded race, the imputed racial distribution is almost identical to the true race distribution, with differences between the true and imputed fraction Minority being less than 0.5% for All Citations, Moving Violations, and Non-Moving Violations, respectively. In other words, my imputation procedure has extremely high internal validity.

257. The third row of numbers in Table 10 shows that among those with unrecorded race, my imputation procedure suggests 64.8% were Minorities overall citations. For Moving Violations this figure is 52.2% Minority and for Non-Moving violations this figure is 70.4% Minority. Comparing this to the above numbers in Table 10, these results reveal that my imputation procedure suggests that a somewhat smaller fraction of citations with unrecorded race went to Minorities than citations with recorded race, *i.e.*, that race information was slightly less likely to be recorded for Non-Minorities than it was for Minorities.

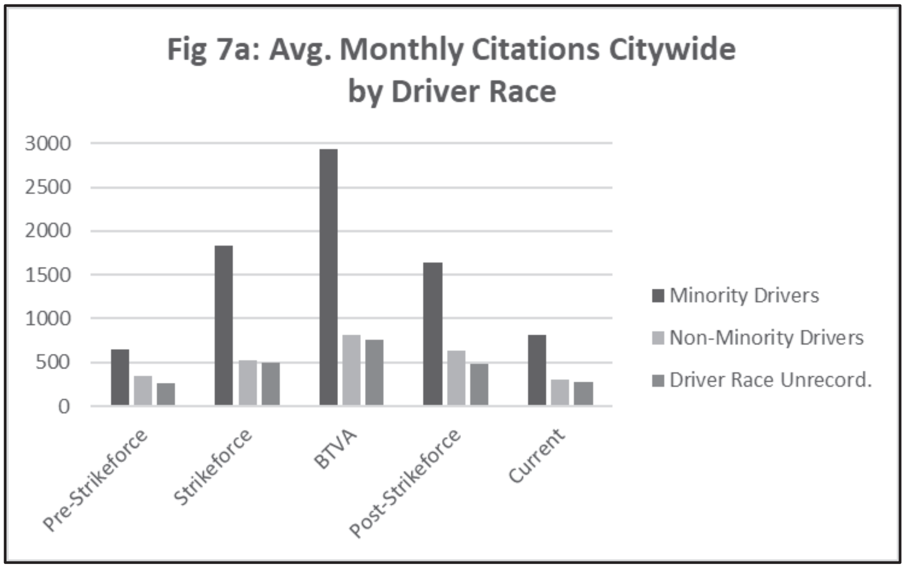
258. The lower panels of Table 10 show the imputed racial distribution by tract type. In general, among those with recorded race, a smaller fraction of citations were issued to Minorities in Low-Minority tracts relative to High-Minority tracts. My imputation procedure also suggests this is true among citations with unrecorded race. It is again also true that my imputation procedure consistently suggests a lower fraction of those citations with unrecorded race to be Minority than those with recorded race across all tract types. Nevertheless, the only category in which Non-Minorities receive more than half the citations is Moving violations in Low-Minority tracts, though almost 40% of those with unrecorded race still seem to be Minorities in Low-Minority tracts, with this fraction being even higher for those with recorded race. This 40% figure for (imputed) Minority tickets in Low-Minority tracts compares to the average minority population percentage in such tracts of 22%.

b) Descriptive Analysis of Citations by Race of Driver

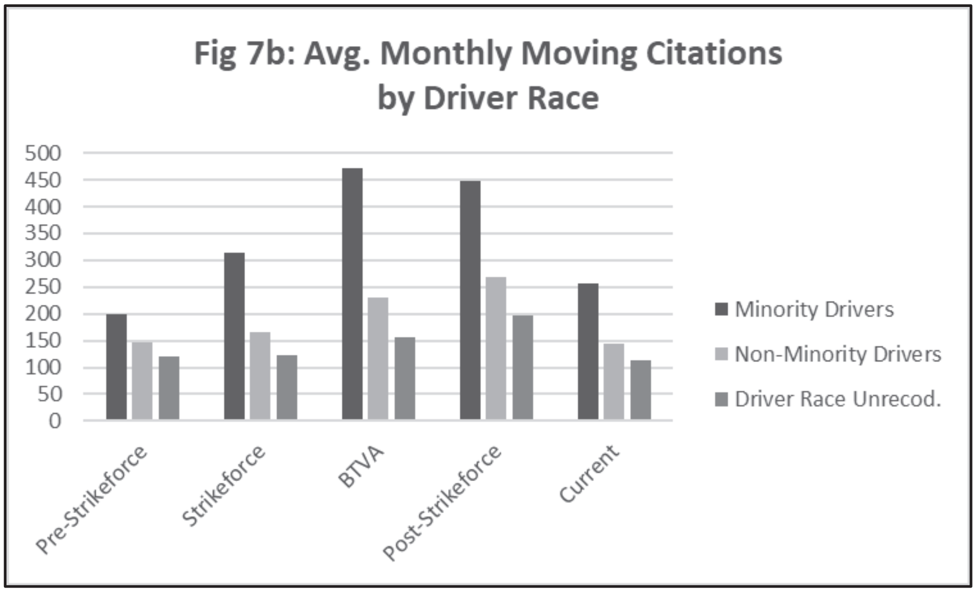
259. Before delving into a more detailed analysis incorporating these imputed racial compositions for citations of unrecorded race, I describe the basic patterns using a relatively simple graphical analysis. For this analysis, I separately analyze those for whom race is recorded from those for who it is not, and I do not use the imputations of the racial distribution of drivers of unrecorded race.

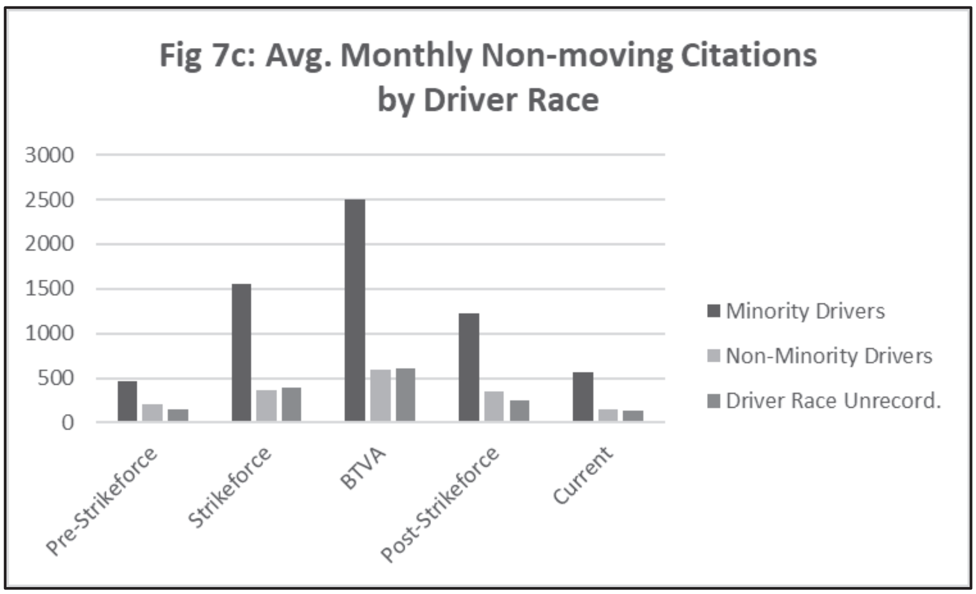
260. Figure 7a shows the average number of monthly citations issued citywide in each era by the recorded driver race. As can be seen, Minority drivers were substantially more likely than Non-Minority drivers to be cited throughout the entire time-frame in question. In the Pre-Strikeforce era, those recorded as Minority drivers citywide received about 650 citations per month, while those recorded as Non-Minority received about 340 citations per month. Moving to the Strikeforce, BTVA, and Post-Strikeforce eras, we can see that citations issued to those recorded as Non-Minority drivers increased modestly to 500–800 per month citywide, while citations issued to those recorded as Minority drivers jumped strikingly to 1,600–3,000 per month during these eras. During the Current era, citations by driver race then returned to roughly similar levels as occurred in the Pre-Strikeforce era, although citation rates for Minority drivers remained somewhat elevated when compared to the Pre-Strikeforce era and still notably higher compared to the rates at which Non-Minority drivers received citations during the Current era. In each era, citations issued to drivers of unrecorded race were generally similar in number to citations issued to Non-

Minority drivers in each era. As discussed above, I estimate that about two-thirds of these individuals with unrecorded race are Minorities.



261. Figures 7b and 7c reflect a similar analysis as was shown in 7a, but separately for Moving violation citations and Non-Moving violation citations. As can be seen, the vast discrepancies in citations by driver race that occurred in the Strikeforce, BTVA, and Post-Strikeforce eras that are seen in Table 7a primarily reflect discrepancies in citations for Non-Moving violations rather than Moving violations (note change of scale on vertical axes).

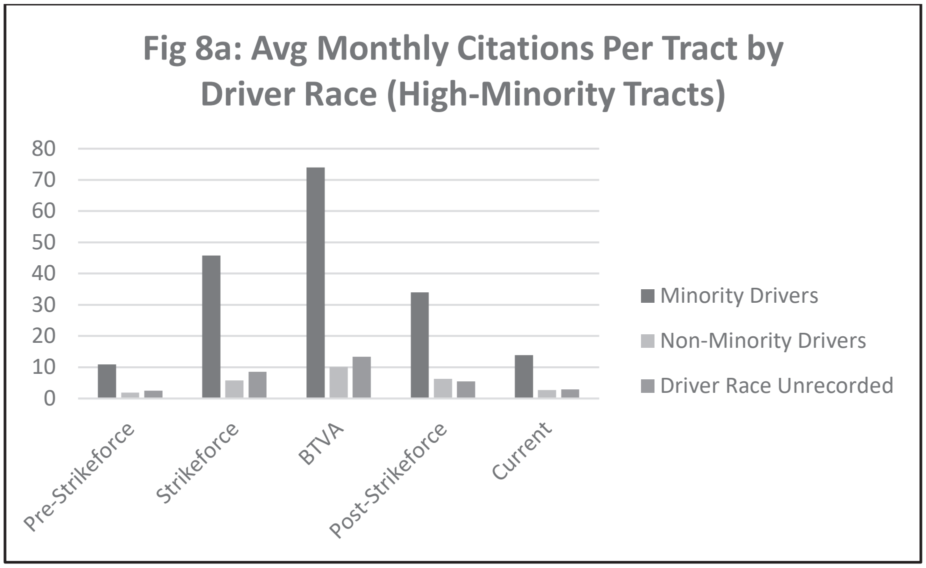


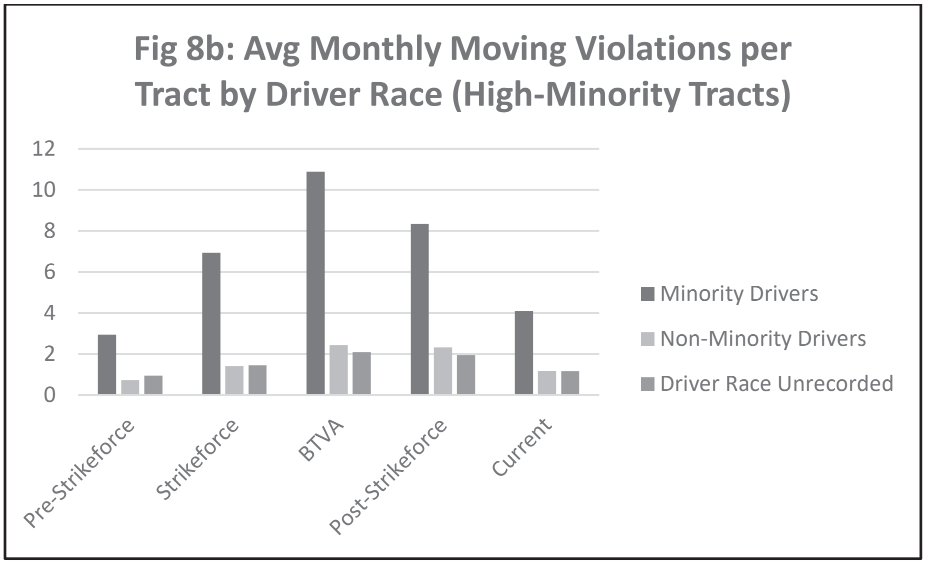


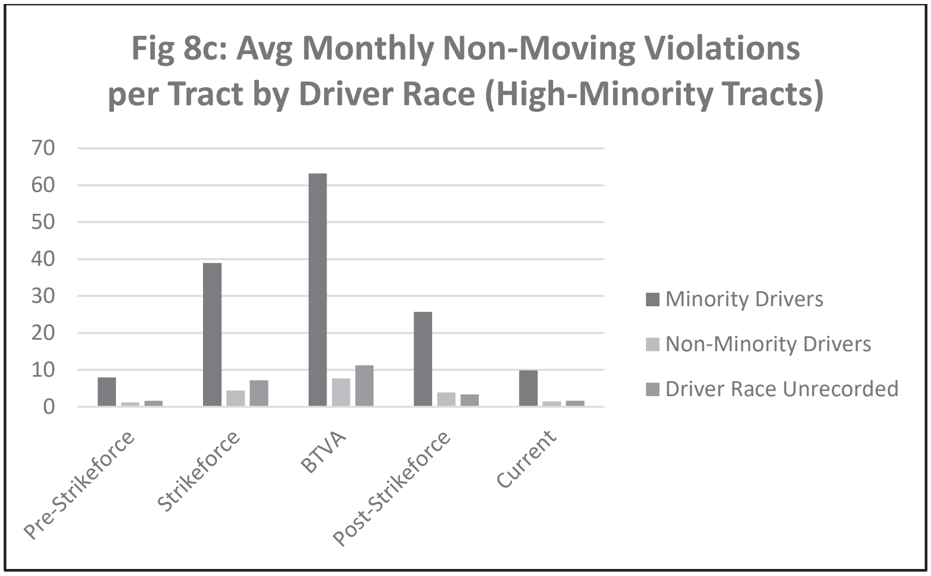
262. As highlighted in the previous sub-section, the vast majority of the increased citation activity that occurred in the Strikeforce, BTVA, and Post-Strikeforce eras took place in High-Minority census tracts, though there were also some modest increases in citations in Mixed-

Race and even Low-Minority tracts during these eras as well. I next look at the racial composition of citations within each of these tract types.

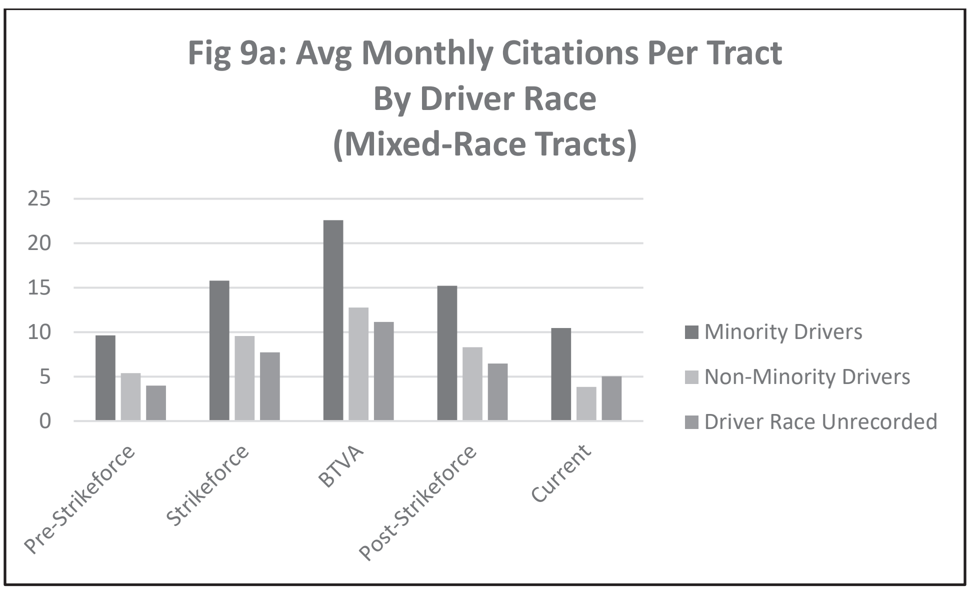
263. Figures 8a–8c look at the racial composition of citations per month in High-Minority tracts across eras. Looking first at Figure 8a, we can see that in High-Minority tracts, very few citations were issued to drivers recorded to be Non-Minority across all five eras. Figures 8b and 8c show that in absolute terms, these racial discrepancies in monthly citations are notably more pronounced for Non-Moving violations than Moving violations (note change in scale of vertical axis).

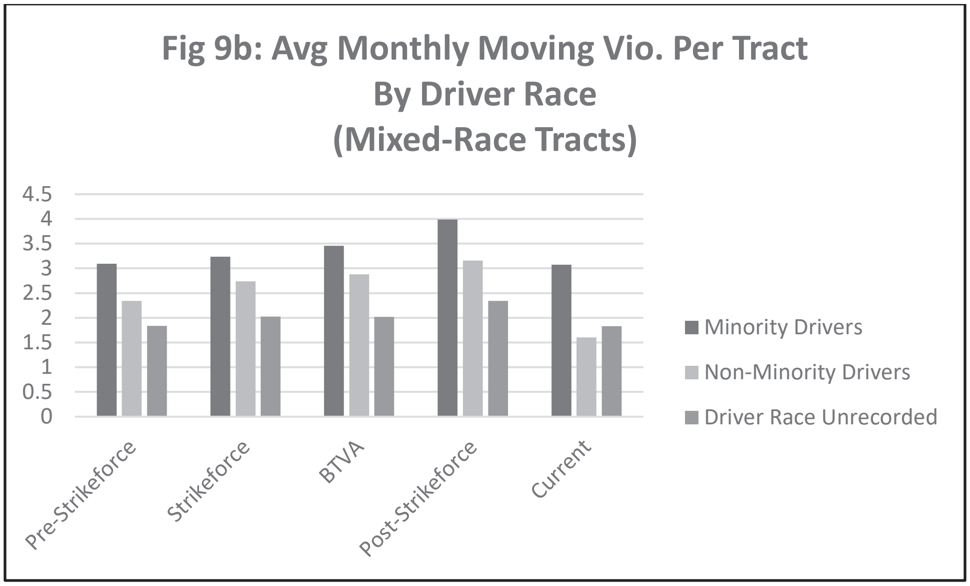


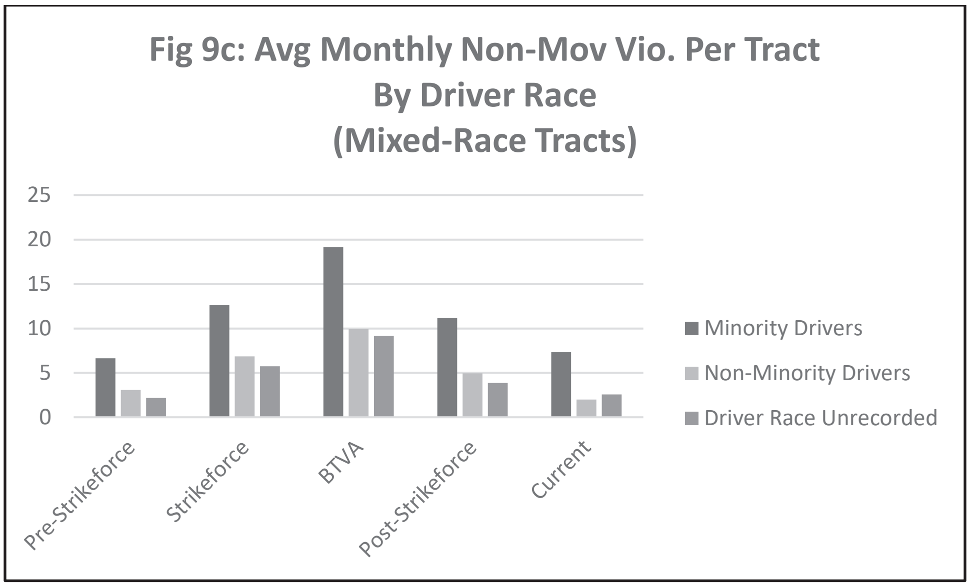




264. Figures 9a–9c look at analogous results for Mixed-Race tracts. As can be seen in Figure 9a, in Mixed-Race tracts, those recorded as Minorities still receive almost twice as many monthly citations as those recorded as Non-Minorities across all five eras. Moreover, as can be seen in Figures 9b and 9c, these racial discrepancies in citations in Mixed-Race tracts are again much more pronounced with respect to Non-Moving violations than Moving violations.

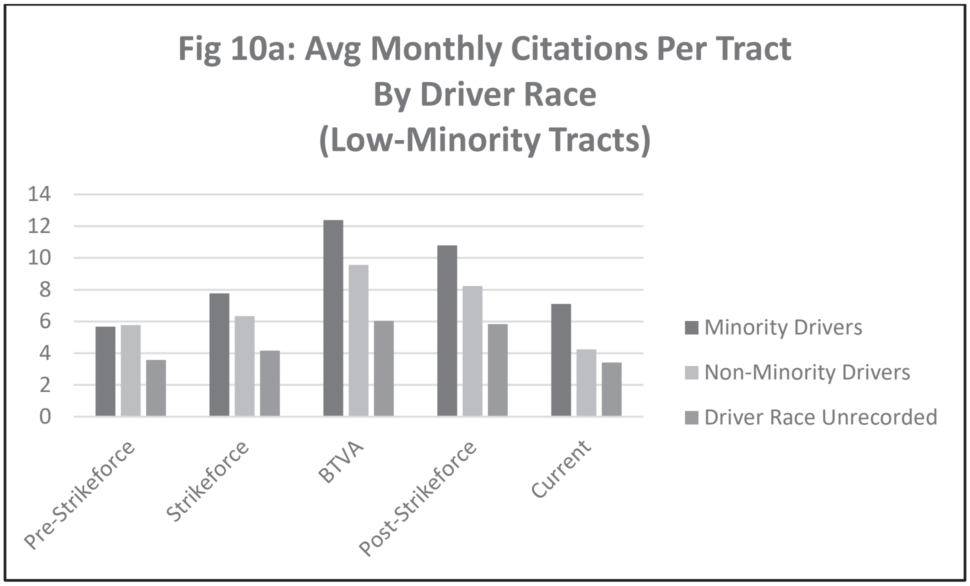


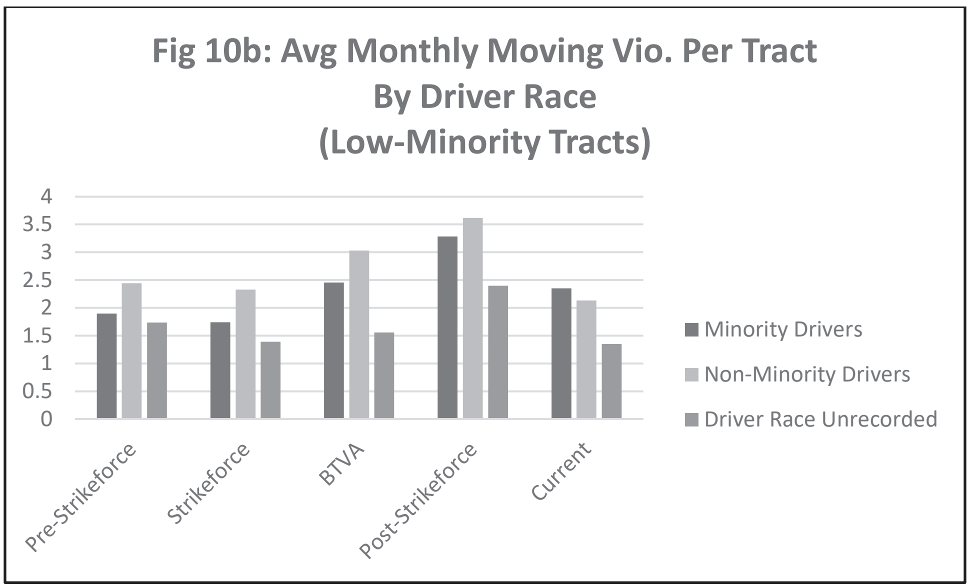


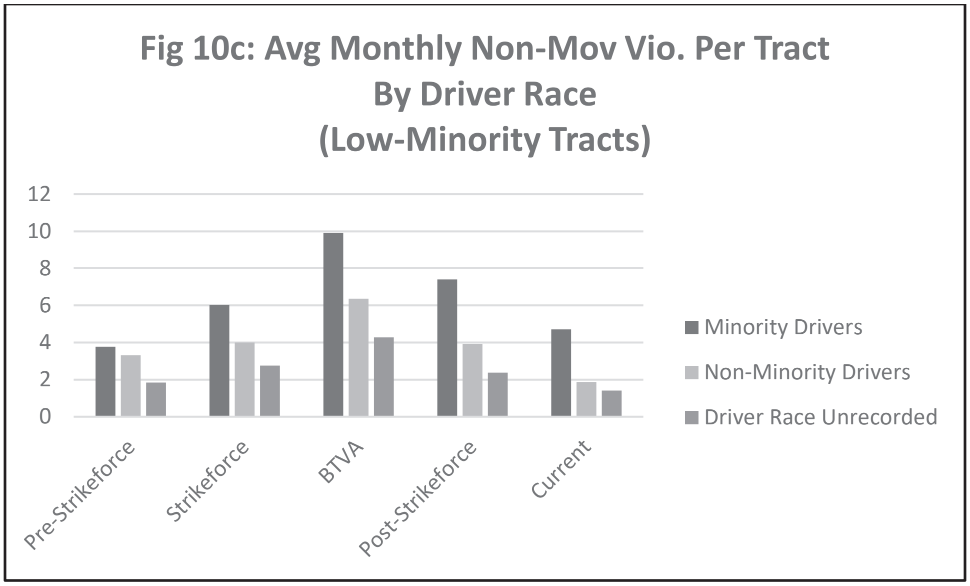


265. Finally, Figures 10a–10c show the analogous results for Low-Minority tracts. As can be seen in Figure 10a, even in Low-Minority tracts, there are generally at least as many monthly citations given to those recorded as Minorities as there are citations given to those recorded as Non-Minorities, even though Minorities by definition make up less than 40% of the residents of these tracts (and, on average, make-up 22% of the residents of these tracts). Moreover, as can be seen in Figures 10b and 10c, while there is little racial discrepancy in Moving violation citations across eras in Low-Minority tracts (with Non-Minorities receiving more Moving violation citations than Minorities in these tracts in each era), there are notably more Non-Moving citations given to those recorded as Minority than those classified as Non-Minority in Low-

Minority tracts in each era. These racial discrepancies are largest in absolute magnitude in the Strikeforce, BTVA, and Post-Strikeforce eras.







266. The race-of-driver analysis discussed in this section and illustrated in Figures 7a–7c, 8a–8c, 9a–9c, and 10a–10c demonstrates that (i) Minority drivers receive significantly more citations on average than Non-Minority drivers; and (ii) this disparity does not simply result from the BPD's issuance of many more citations in High-Minority tracts than Low-Minority tracts over substantial portions of the time period examined here. Indeed, in Mixed-Race and even Low-Minority tracts, those recorded as Minorities received more citations than those recorded as Non-Minorities.

267. Finally, one can assess the extent to which any of the above conclusions would be affected by adding in the citations for drivers of unrecorded race. Recall the imputed racial distributions of tickets issued to drivers of unrecorded race in Table 10. If one were to apply those percentages to distribute the race-unrecorded citations to the Minority Drivers and Non-Minority Drivers bars in the above figures, it is clear that none of the above conclusions would change, as even among tickets issued in Low-Minority tracts, the imputation procedure suggests that almost half went to Minorities. Only if one assumes that almost all of those drivers for whom race is unrecorded are Non-Minority—an assumption that is almost certainly incorrect given the cross-validation and imputation procedures discussed earlier—would it no longer be true that Minorities are issued the majority of citations even in Low-Minority tracts.

c) *Analysis of Racial Disparities in Citations Across Individuals Within Neighborhoods*

268. This section assesses racial discrepancies in ticketing within neighborhoods in a more detailed manner. As described in the Methodology section, using census tracts as my definition of neighborhoods, I can compute Population Adjusted Ticketing Rates for each race in each tract for each era. For each tract, this metric measures the average monthly tickets issued to members of each race for every 100 individuals of that race residing in that tract. To assess racial disparity within tracts overall across the city, I average the Population Adjusted Ticketing rates for each race across all tracts, weighted by total tract population. I then use a t-test to analyze whether average Population Adjusted Ticketing rates within tracts differ by race in the City of Buffalo as whole. As discussed in the Methodology section, I include those citations issued to drivers of unrecorded race using the imputation procedure I laid out earlier.[31]

269. Table 11 shows the results of this exercise by era for all citations, as well as analyzing Moving and Non-Moving violation citations separately. The first two columns of numbers show average tract-level Population Adjusted Ticketing rates for each race (weighted by total tract population), the third column of numbers shows the difference in these measures across races, the third and fourth columns of numbers show the test statistic and associated p-value, and the final column shows the ratio of average Minority to Non-Minority Population Adjusted Ticketing rates.

[31] Given that my imputation procedure generally implies a smaller fraction of drivers with unrecorded race were Minorities relative to drivers with recorded race, excluding drivers of unrecorded race from this exercise would suggest even *greater* racial disparity than I show below.

Table 11 - Population Adjusted Monthly Ticketing Rates by Race, Avg within Tracts						
	Pop. Adj. Monthly Ticketing Rates (Within Tract Pop. Wgt. Avg.)					Min. to Non-Min.
Era	**Minorities**	**Non-Minorities**	**Diff**	**t-stat**	**pval**	**Ratio**
All Citations						
Pre-Strikeforce	0.95	0.40	0.55	5.1	0.000	2.4
	(0.11)	(0.04)	(0.11)			
Strikeforce	1.69	0.96	0.73	5.0	0.000	1.8
	(0.17)	(0.16)	(0.15)			
BTVA	2.51	1.79	0.71	2.9	0.005	1.4
	(0.25)	(0.37)	(0.24)			
Post-Strikeforce	1.75	1.00	0.75	4.1	0.000	1.8
	(0.18)	(0.16)	(0.18)			
Current	0.95	0.45	0.51	4.4	0.000	2.1
	(0.11)	(0.08)	(0.12)			
Moving Violations						
Pre-Strikeforce	0.30	0.18	0.12	3.6	0.001	1.7
	(0.04)	(0.02)	(0.03)			
Strikeforce	0.33	0.25	0.08	2.8	0.006	1.3
	(0.03)	(0.04)	(0.03)			
BTVA	0.44	0.39	0.05	1.0	0.315	1.1
	(0.04)	(0.06)	(0.05)			
Post-Strikeforce	0.51	0.39	0.12	1.9	0.056	1.3
	(0.05)	(0.06)	(0.06)			
Current	0.32	0.21	0.10	2.1	0.037	1.5
	(0.04)	(0.04)	(0.05)			
Non-Moving Violations						
Pre-Strikeforce	0.65	0.23	0.42	5.3	0.000	2.9
	(0.09)	(0.02)	(0.08)			
Strikeforce	1.36	0.71	0.65	5.2	0.000	1.9
	(0.15)	(0.13)	(0.12)			
BTVA	2.05	1.39	0.65	3.2	0.002	1.5
	(0.22)	(0.31)	(0.2)			
Post-Strikeforce	1.15	0.57	0.58	5.5	0.000	2.0
	(0.11)	(0.11)	(0.11)			
Current	0.60	0.21	0.39	5.4	0.000	2.9
	(0.08)	(0.04)	(0.07)			

*Standard errors in parentheses, adjusted for multiple observations per tract in eras with more than one six-month period. See text regarding how tickets issued to driver's with unrecorded race were allocated across race groups. For each racial group r, Population Adjusted Ticketing Rates measure the average number of tickets issued to members of racial group r per month in each tract per hundred people of race r in that tract (see text for details).

270. Looking first at the results for all citations in the top panel of Table 11, we see that in all eras, average within-neighborhood Population Adjusted Ticketing rates are statistically significantly higher for Minorities than Non-Minorities at the 1% level (indeed at the 0.1% level). As the last column shows, the population-weighted average of these within-tract Population Adjusted Ticketing rates for Minorities are 1.4–2.3 times higher than they are for Non-Minorities. Interestingly, these ratios are highest in the Pre-Strikeforce era and the Current era—eras in which High-Minority and Low-Minority tracts were ticketed at roughly similar levels of intensity. I will return to discuss this finding in more detail below.

271. The lower panels of Table 11 show the results separately for Moving violations and Non-Moving violations. As can be seen, the disparities in average within- tract Population Adjusted Ticketing rates are much larger with respect to Non-Moving violations than Moving violations. In other words, Minorities were not only more likely than Non-Minorities to be ticketed within any given tract, but they were also much more likely than Non-Minorities to receive Non-Moving violation tickets within that tract. Indeed, in only two of the five eras (the Pre-Strikeforce era and Strikeforce era) were the within-tract Population Adjusted Ticketing rates for Moving violations statistically different between Minorities and Non-Minorities at the 1% level. By contrast, when it comes to Non-Moving violations, average within-tract Population Adjusted Ticketing rates differed significantly between Minorities and Non-Minorities at well below the 1% level throughout the entire period analyzed. Moreover, the ratio of Minority to Non-Minority within-tract Population Adjusted Ticketing rates for Non-Moving violations exceeded 1.5 in all eras. Indeed, it exceeded 2.5 in two of the eras.

272. Overall, these results indicate that there were substantial racial disparities in tickets relative to population within neighborhoods (as defined by census tracts) across all eras, particularly when it came to Non-Moving violations. As compared to some of the across-neighborhood findings discussed above, these disparities did not moderate or fade to nearly the same extent in the Current Era: Even in the Current Era, Minority drivers are over two and a half times as likely as Non-Minority individuals to be ticketed within a particular neighborhood.

3. Summary and Decomposition of Racial Disparities in Citations Within and Across Neighborhoods

273. The previous sub-section analyzed racial disparities in citation rates *within* neighborhoods as defined by census tracts, or how the number of citations per resident differs by

race within tracts. By contrast, earlier results examined racial disparities in citations *across* neighborhoods as defined by census tracts, or how citations differ in tracts that are primarily Minority relative to tracts that are primarily Non-Minority. This subsection analyzes (i) how the number of citations per resident differs by race citywide, and (ii) how much of this citywide racial disparity in citations per resident resulted from disparities in citations *within* neighborhoods (*i.e.*, from disproportionate ticketing of Minority drivers within a given tract) versus *across* neighborhoods (*i.e.*, from disproportionate issuance of tickets in tracts with larger Minority populations).

274. To do so, I calculate citywide Population Adjusted Ticketing rates for Minorities and for Non-Minorities for each era. The methodology for this statistic is the same as described in the Methodology section, except that here I conduct this analysis for the City as a whole rather than separately by census tract. The first three columns of numbers in Table 12 show the results of this exercise.

Table 12 - Decomposition of Racial Differences in Population Adjusted Ticketing Rates, Within vs Across Tract Disparities								
	Population Adjusted Monthly Ticketing Rates (Citywide)			Population Adjusted Monthly Ticketing Rates (Within Tract Pop. Wgt. Avg.)			% of Overall Racial Disparity in Ticketing Rates Due to	
Era	Minorities	Non-Minorities	Citywide Ratio	Minorities	Non-Minorities	Within Tracts Ratio	Within Tract Disparities	Across Tract Disparities
All Citations								
Pre-Strikeforce	0.64	0.33	**1.9**	0.95	0.40	**2.4**	148%	-48%
Strikeforce	1.77	0.50	**3.6**	1.69	0.96	**1.8**	30%	70%
BTVA	2.80	0.78	**3.6**	2.51	1.79	**1.4**	15%	85%
Post-Strikeforce	1.57	0.61	**2.6**	1.75	1.00	**1.8**	48%	52%
Current	0.78	0.29	**2.7**	0.95	0.45	**2.1**	66%	34%
Moving Violations								
Pre-Strikeforce	0.20	0.15	**1.3**	0.30	0.18	**1.7**	221%	-121%
Strikeforce	0.30	0.16	**1.9**	0.33	0.25	**1.3**	37%	63%
BTVA	0.44	0.21	**2.0**	0.44	0.39	**1.1**	13%	87%
Post-Strikeforce	0.43	0.26	**1.6**	0.51	0.39	**1.3**	48%	52%
Current	0.24	0.14	**1.8**	0.32	0.21	**1.5**	64%	36%
Non-Moving Violations								
Pre-Strikeforce	0.44	0.18	**2.4**	0.65	0.23	**2.9**	132%	-32%
Strikeforce	1.47	0.34	**4.3**	1.36	0.71	**1.9**	27%	73%
BTVA	2.36	0.56	**4.2**	2.05	1.39	**1.5**	15%	85%
Post-Strikeforce	1.11	0.31	**3.6**	1.15	0.57	**2.0**	61%	39%
Current	0.52	0.13	**3.9**	0.60	0.21	**2.9**	64%	36%

*See text regarding how tickets issued to driver's with unrecorded race were allocated across race groups. For each racial group r, Population Adjusted Ticketing Rates measure the average number of tickets issued to members of racial group r per six months in each tract per hundred people of race r in that tract (see text for details).

275. There are several findings of interest in the first three columns of numbers in Table 12. First, the large increase in ticketing that began with the implementation of the Strikeforce and continued through the Post-Strikeforce era documented earlier in this report substantially increased citywide Population Adjusted Ticketing rates for Minorities relative to Non-Minorities. Indeed, the ratio of Minority to Non-Minority citywide Population Adjusted Ticketing rates goes from under 2 in the Pre-Strikeforce era to over 3.5 in the Strikeforce and BTVA eras. These ratios stay

above 2.5 through the Current era. In other words, citywide, Minority drivers went from receiving just under twice as many citations per capita relative to Non-Minorities in the Pre-Strikeforce era, to over three and a half times as many citations per capita in the Strikeforce and BTVA eras. Following the disbanding of the Strikeforce, Minority drivers still continued to receive around two and a half times as many citation as Non-Minorities through the Current era.

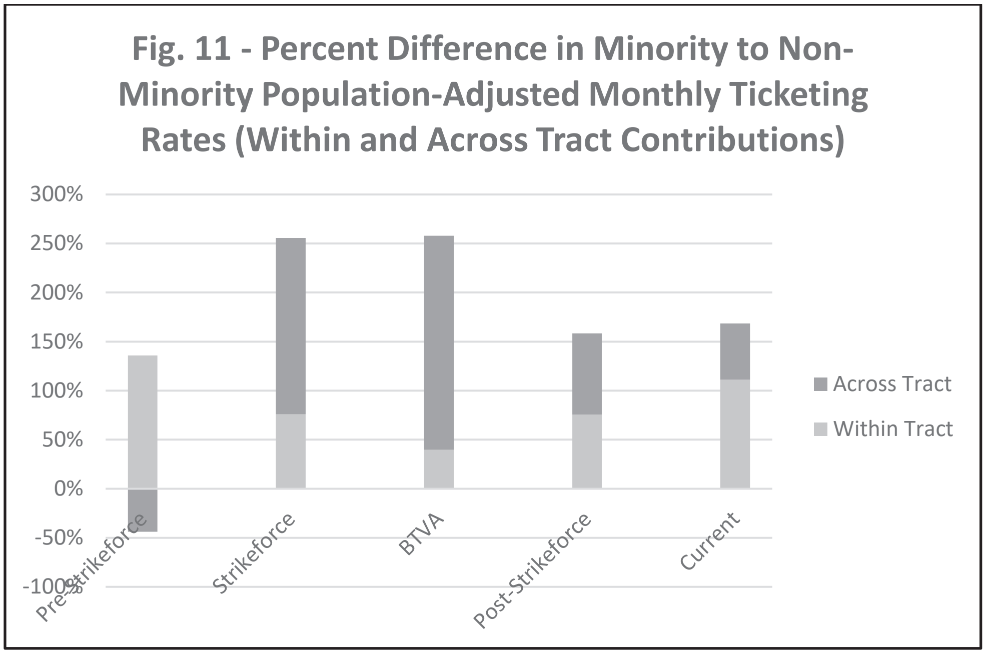
276. Looking at the first three columns of numbers in the lower panels of Table 12, we again see that at the citywide level, racial differences in Population Adjusted Ticketing rates are generally twice as large for Non-Moving violations relative to Moving violations for each era. Notably, with respect to Non-Moving violations, citywide Population Adjusted Ticketing rates are well over 4 times higher for Minorities than Non-Minorities during the Strikeforce and BTVA eras.

277. The fourth through sixth columns of numbers in Table 12 simply re-show the results from Table 11 in a slimmed down fashion. What this helps reveal is how within- tract racial disparities in Population Adjusted Ticketing rates compare to overall citywide disparities in Population Adjusted Ticketing. One can then decompose the overall citywide racial disparity in Population Adjusted Ticketing rates into what fraction arises due to within-tract disparity and what fraction arises due to across-tract disparities. To do so, I take the within-tracts ratio of Minority to Non-Minority average Population Adjusted Ticketing Rates and subtract one from it. This is a measure of the "excess" ticketing of Minorities on average within tracts. I then do the same using the citywide ratio, which is a measure of the "excess" ticketing of Minorities that arises both *within* and *across* tracts. If I then divide the former by the latter, I get the fraction of overall excess ticketing of Minorities that arises within tracts. Subtracting this from one then gives the fraction of overall excess ticketing of Minorities that arises across tracts.

278. The last two columns of numbers in Table 12 show the results of this decomposition exercise by era for all citations, as well as for Moving and Non-Moving violations separately. To be clear, these last two columns simply show what *fraction* of the citywide racial disparity in Population Adjusted Ticketing rates shown in the third column of numbers is attributable to *within*-tract relative to *across*-tract disparities.

279. Figure 11 summarizes the main findings from Table 12. The height of the bars shows the percent difference in Population Adjusted Monthly Ticketing rates between Minorities and Non-Minorities overall. The shading indicates how much of this overall disparity in each era

arises due to within-tract disparities (light grey) and how much of the overall disparity arises due to across-tract disparities (dark grey).



280. Looking at the first bar on the left in Figure 11 (which reflects the top row of Table 12), we see that during the Pre-Strikeforce era, across-tract disparities actually offset some of the within-tract disparities. Specifically, while within-tract Population Adjusted Monthly Ticketing rates were close to 140% higher among Minorities than Non-Minorities, slightly lower ticketing rates in the higher Minority tracts relative to lower Minority tracts during this era meant that *overall* Population Adjusted Monthly Ticketing rates were about 90% higher among Minorities than Non-Minorities. As shown by the rightmost two columns in Table 12, within-tract disparities account for 148% of overall disparity in the Pre-Strikeforce era, while across-tract disparities account for -48% of the overall disparity. In other words (and as documented above), while Minorities received citations at substantially higher rates relative to their populations in most neighborhoods during the Pre-Strikeforce era, higher Minority neighborhoods were actually receiving a slightly lower number of tickets relative to their populations. While these across-tract disparities partially offset the within-tract disparities, Minorities still received about 90% more (or close to twice as many) tickets relative to their population citywide during the Pre-Strikeforce era on net, compared to Non-Minorities.

281. This pattern changed dramatically during the Strikeforce era. As can be seen in the second bar from the left in Figure 11 (and can be inferred from the third column of numbers in the

second row of Table 12), overall Population Adjusted Monthly Ticketing rates were more than 250% higher among Minorities than Non-Minorities during the Strikeforce era. As can be seen in Figure 11 (and the last two columns in in the second row Table 12), these overall racial disparities in Population Adjusted citation rates reflected higher citation rates among Minorities relative to Non-Minorities both *within* and *across* tracts. However, across-tract disparities accounted for over two-thirds of the overall difference in Population Adjusted Monthly Ticketing rates across races during this time era. Thus, although Minority drivers continued to receive tickets at greater rates than Non-Minority drivers within neighborhoods, they now *also* received even more citations, on average, because BPD began issuing a greater number (and proportion) of tickets in neighborhoods with larger Minority populations.

282. Looking at the third bar in Figure 11 (which reflects data from the third row of Table 12), we can see a similar finding for the BTVA era. Overall, Population Adjusted Citation rates were about 250% higher for Minorities relative to Non-Minorities, which reflected both within-tract and across-tract disparities. As in the Strikeforce era, however, well over two-thirds of the overall racial disparity in Population Adjusted Citation rates arose due to across-tract disparities. These findings for the Strikeforce and BTVA eras are unsurprising given the findings, discussed earlier in this report, that during the Strikeforce and BTVA eras, the number of citations issued High-Minority tracts increased to double or even triple the number of citations issued in Low-Minority tracts. This large jump in ticketing in the higher Minority neighborhoods would account for why across-tract racial disparities in Population Adjusted Citation rates account for a much higher portion of the racial disparities in Citywide Population Adjusted Citation rates for these two eras.

283. As can be seen in looking at the Post-Strikeforce and Current era results as summarized in the final two bars of Figure 11 (which reflect the fourth and fifth rows of Table 12), these patterns changed again after the Strikeforce was disbanded. In both the Post-Strikeforce and Current eras, overall Population Adjusted Citation rates were about 150% higher among Minorities than Non-Minorities. While still large, such racial disparities in citations were modestly smaller than they were during the Strikeforce and BTVA eras. Moreover, during the Post-Strikeforce and Current eras, about half to two-thirds of the overall racial disparities in citation rates arose *within* tracts relative to *across* tracts. This is unsurprising given the earlier findings that tract-level racial

disparities in ticketing between High-Minority and Low-Minority tracts fell following the dissolution of the Strikeforce and mostly dissipated by the Current era.

284. In summary, Table 12 and Figure 11 show that there were substantial racial disparities in ticketing per capita over all the eras examined here, including the Current era, with these disparities being particularly large during the Strikeforce and BTVA eras and generally much larger with respect to Non-Moving violations than Moving violations. Using census tracts as my definition of neighborhoods, I find that in the Pre-Strikeforce, Post-Strikeforce, and Current eras, these racial disparities in per capita ticketing primarily arose *within* neighborhoods rather than due to differential ticketing *across* neighborhoods. However, during the Strikeforce and BTVA eras, racial disparities in per capita ticketing were dominated by differential ticketing *across* neighborhoods rather than disparities *within* neighborhoods. This reflected the large volume of citations issued in higher Minority neighborhoods relative to lower Minority neighborhoods during these two eras.

4. Analysis of Multiple Citations per Incident by Race of Driver

285. As described in the Data section above, I was able to collapse the citation data to the incident level by bundling all citations that were issued at the same location on the same hour/day/minute into a particular incident.[32] This subsection uses this incident-level data to examine a particular type of police practice, namely the issuance of more than one citation in a given incident, by race of driver.

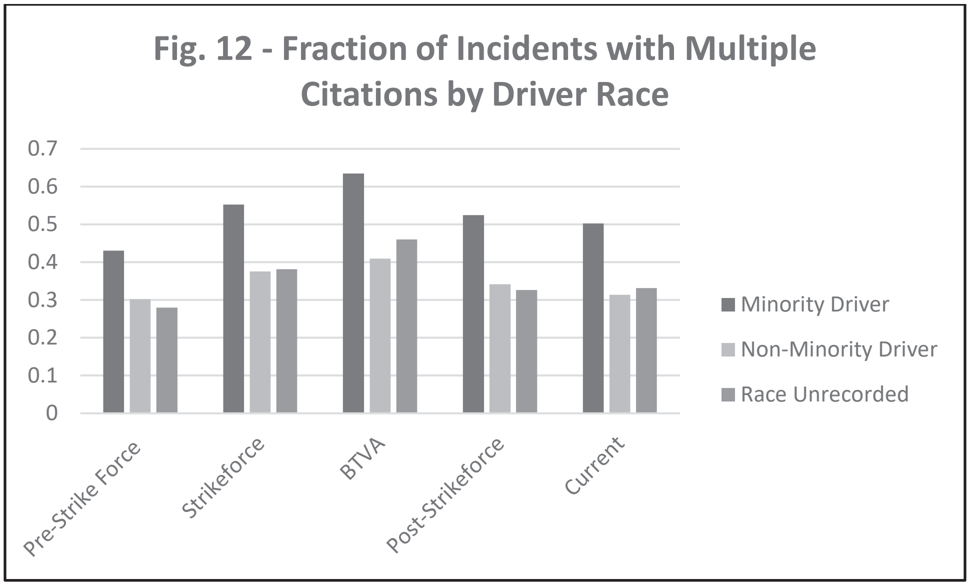
286. Among those incidents for which race was recorded, 55% of those incident involved the issuance of one or more tickets to Minorities. As documented earlier, however, among citations with recorded race, 76% of citations were issued to Minorities. This implies that incidents involving Minorities were much more likely to involve multiple citations than incidents involving Non-Minorities.

287. Further analysis confirms that this is true: Minorities were more likely than Non-Minorities to receive multiple citations in single incidents. Multiple citations were issued in 37% of incidents involving a driver categorized as Non-Minority. By contrast, multiple citations were

[32] There were also a relatively small number of cases where I bundled together tickets as being in the same incident when they occurred at the same day/hour/minute and to the same person as indicated by their license number/license reference number, but at not identically written locations. This was usually due to either typos in the location, or spelling out "Avenue" vs. "Ave", or "North" vs. "N".

issued in 56% of incidents involving a driver categorized as Minority. Multiple citations were issued in 40% of incidents with unrecorded driver race.

288. Figure 12 shows the fraction of incidents in which multiple citations were issued by race across eras. As can be seen, racial disparities in the fraction of incidents in which multiple citations were issued were quite consistent across all eras. Indeed, from the Strikeforce era onward, approximately 50–65% of incidents involving Minority drivers resulted in multiple citations, while just 30–40% of incidents involving Non-Minority drivers resulted in multiple citations.



289. Figure 12 also shows that the likelihood of multiple tickets within a given incident for those with unrecorded race generally lies at or slightly above the analogous likelihood for those recorded as Non-Minorities, but generally well below the analogous likelihood for those recorded as Minorities. This might seem a little surprising given that my interpolation procedure suggested that roughly 65% of citations issued to drivers of unrecorded race were Minority. However, note that Figure 12 is at the incident level. If Minorities with unrecorded race are also more likely to receive multiple tickets per incident than Non-Minorities of unrecorded race, then it follows that the percentage of incidents with unrecorded race that actually involved Minorities will be smaller than the percentage of citations with unrecorded race that were actually issued to Minorities.

290. As described in the Methods section above, to look in more depth at racial disparities in the frequency with which BPD issued multiple citations per incident, I first regress an indicator for whether or not multiple citations were issued in each incident on indicator variables for the time the incident occurred, as well as driver age, using just the sample of incidents that

involved Non-Minorities. I do this for each era separately. I then predict the likelihood of multiple citations for each incident using the coefficients obtained from the previous regressions and the time-of-day indicators and age of driver for each incident. I then calculate the difference between the actual incidence of multiple citations and the predicted likelihood of multiple citations for each incident and assess whether this difference is significantly different from zero in incidents involving Minorities and incidents in which race is unrecorded. The results of this exercise are shown in Table 13.

Table 13 - Difference Between Actual and Predicted Likelihood of Multiple Citations within Incident					
	Coeff. On Predictors of Multiple Citations for Non-minority Drivers (Probit)				
Variable	Pre-Strikeforce Era	Strikeforce Era	BTVA Era	Post-Strikeforce Era	Current Era
Age	-0.004	-0.004***	-0.007***	-0.002	-0.004***
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
12PM-6PM	0.293**	0.334***	0.466***	0.304***	0.219***
	(0.127)	(0.057)	(0.077)	(0.072)	(0.057)
6PM-12AM	0.523***	0.475***	0.616***	0.594***	0.298***
	(0.116)	(0.061)	(0.091)	(0.094)	(0.083)
12AM-6AM	0.638***	0.534***	0.410***	0.525***	0.587***
	(0.141)	(0.068)	(0.094)	(0.106)	(0.094)
Constant	-0.793***	-0.566***	-0.391***	-0.639***	-0.546***
	(0.147)	(0.065)	(0.082)	(0.092)	(0.093)
Observations	1,440	11,101	11,528	8,620	5,805
	Mean Difference Between Actual and Predicted Likelihood of Multiple Citations w/in Incident				
Minority Drivers	0.124***	0.169***	0.189***	0.145***	0.164***
	(0.015)	(0.008)	(0.012)	(0.015)	(0.016)
Driver Race Unrecorded	-0.023	0.002	0.034***	0.008	0.042**
	(0.019)	(0.007)	(0.011)	(0.014)	(0.017)
Intercept (Non-minority Drivers)	0.000	0.000	0.000	0.000	0.000
	(0.013)	(0.010)	(0.018)	(0.017)	(0.020)
*"Excluded" Comparison Time Category 6AM-12PM. All standard errors are heteroskedasticity robust and clustered by census tract. ***indicates significance at 1% level, **indicates significance at 5% level, *indicates significance at 10% level.					

291. The top panel of Table 13 shows the results from the first-stage prediction regressions by era.[33] As can be seen, for incidents involving Non-Minorities, multiple citations per incident are generally more likely during some times of day than others. Consistently fewer incidents involve multiple ticketing when they occur in the morning from 6AM to 12PM (the "excluded" comparison category in Table 13).

292. The bottom panel of Table 13 shows how the actual likelihood of receiving multiple citations per incident differs from what would be predicted by the time of the incident and the age

[33] Given that the dependent variable is binary, I use a Probit regression specification rather than OLS. The results are essentially the same if an OLS specification is used.

of the driver for those classified as Minority (and those for whom race is unrecorded). As can be seen, those classified as Minorities are between 12–20 percentage points more likely to receive multiple citations within a given incident than would be predicted if the rate at which they received multiple tickets per incident was similar to that of the rate at which Non-Minorities of similar age received multiple tickets per incident during incidents occurring at the same time of day. These differences are all statistically significant at the 1% level. The results for incidents with driver race unrecorded are shown in the second to last row. These coefficients are generally quite close to zero, suggesting that those of unrecorded race on average receive multiple tickets within a given incident at roughly the same rate as Non-Minority drivers stopped at the same time of the day. Finally, the last row shows that for Non-Minority drivers, there is on average no difference between the true incidence of multiple citations and the predicted incidence, which should again be true by construction of this method.

293. As discussed above in the Methodology section, incidents that involve tinted windows are a particularly interesting class of incidents when evaluating multiple citations per incident, as almost every car that is eligible to receive one citation for tinted windows should be eligible for additional tinted window citations, as it would seem unlikely that any driver would only tint one window, much less that there would any racial differences in the rates at which drivers tint only a single window.

294. Before proceeding, it is important to note that tinted-window incidents are not uncommon: more than 22,000 such incidents occurred in the City of Buffalo over the time period examined here. Such incidents are far more likely to involve Minorities than Non-Minorities and far more likely to occur in High-Minority or Mixed-Race tracts than Low-Minority tracts. Figures 13a–13c highlight these differences. As can be seen in Figure 13a, there were five and a half times as many tinted window incidents involving individuals recorded as Minorities relative to individuals recorded as Non-Minorities. As can be seen in Figure 13b, if we use the imputations of the racial distribution for those of unrecorded race from Table 10, this discrepancy becomes even larger, with Minorities accounting for 83% of tinted window incidents. Finally, Figure 13c shows the average number of such incidents per tract by tract type. As can be seen, on average, each Mixed-Race tract saw almost twice as many tinted window incidents than each Low-Minority tract, and each High-Minority tract saw almost four times as many such incidents than each Low-Minority tract.

Fig. 13a - Number of Tinted Window Incidents by Race

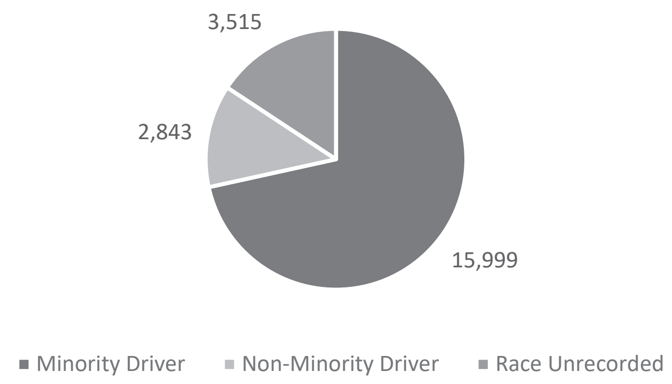


Fig. 13b - Number of Tinted Window Incidents by Race (Including Imputation for Unrecorded Race)

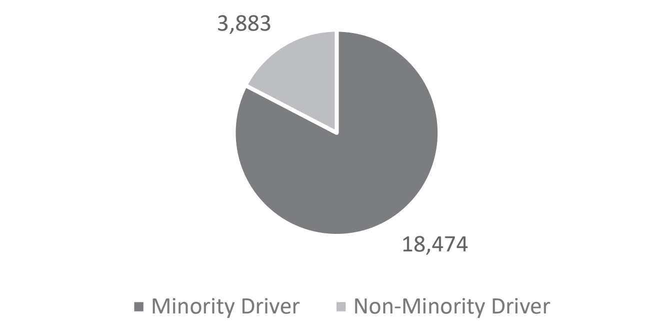
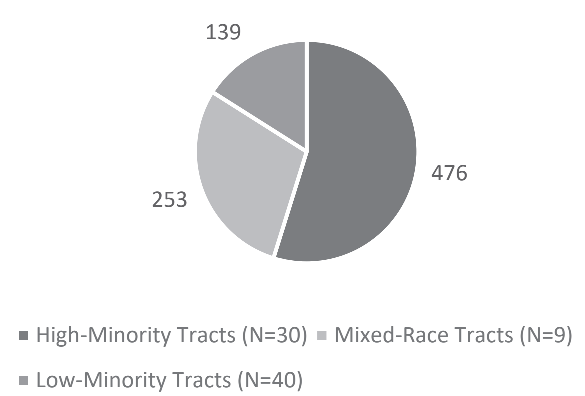
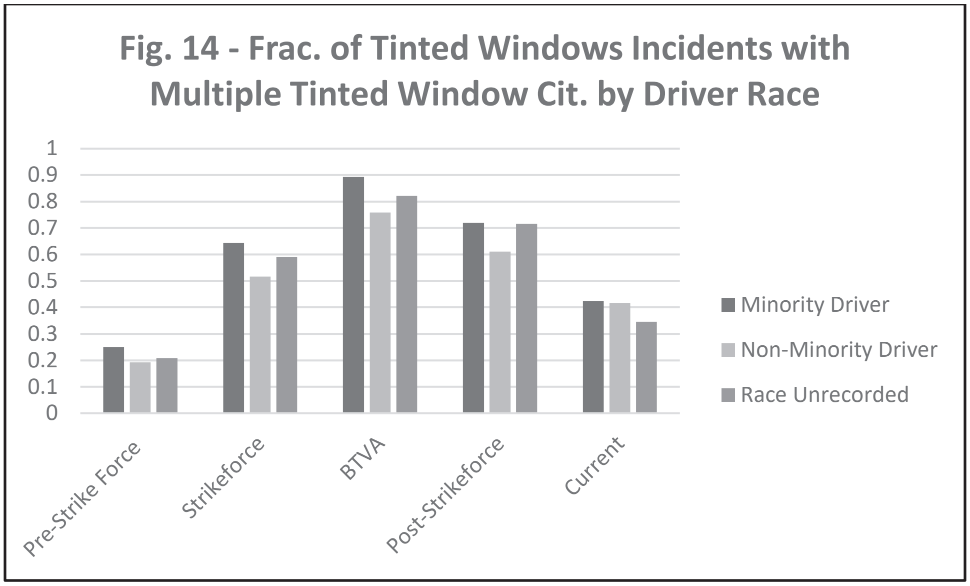


Fig 13c - Avg Number of Tinted Window Incidents per Tract, by Tract Type



295. Figure 14 shows the fraction of tinted window incidents in which multiple tinted windows citations are issued, by driver race across eras. As can be seen, in the Pre-Strikeforce era, a relatively small fraction of tinted window incidents resulted in multiple tinted window tickets across all races. However, starting in the Strikeforce era, the fraction of tinted window incidents that led to multiple tinted widow tickets went up to just over 50% for those recorded as Non-Minorities, but to almost 65% for those recorded as Minorities. During the BTVA era, the rate at which multiple tinted windows tickets were issued in incidents involving tinted windows climbed to 75% for drivers recorded as Non-Minorities but almost 90% for drivers recorded as Minorities during the BTVA era. Rates of multiple tinted window tickets in tinted window incidents remained close to as high in the Post-Strikeforce era, as did the racial discrepancy in such multiple ticketing for tinted windows. Note again that these racial discrepancies in the rates at which drivers were issued multiple tinted window citations in incidents involving at least one tinted window citation are compounding upon the already vast underlying racial discrepancies, highlighted earlier, in the rates at which drivers are cited for tinted windows. In other words, Minority drivers are not only much more likely to be cited for tinted windows in the first place, but also more likely to receive multiple tickets if they are cited for tinted-windows violations. And while it is possible that the underlying rate at which drivers of different races receive tinted-windows citations could reflect underlying differences in the rates at which drivers of different races have their windows tinted (or excessively tinted), it seems unlikely that there are racial differences in the rates at which drivers who have one or more windows tinted have *multiple* windows ticketed, or that underlying behavior with respect to tinted windows changed dramatically between 2012 and 2016.



296. The frequency of this phenomenon diminished in the Current era for drivers of both races, though those recorded as Minorities remain somewhat more likely than those recorded as Non-Minorities to receive multiple tinted window tickets in tinted window incidents.

297. Figure 14 also shows that the likelihood of multiple tinted window tickets in tinted window incidents for those with unrecorded race generally lies a little more than halfway between the analogous likelihood for those recorded as Minorities and those recorded as Non-Minorities. This is consistent with the imputation exercise, which suggests that somewhat more than half of those of unrecorded race are Minority.

298. Table 14 presents regression prediction results analogous to those presented in Table 13 but now limited to the sample of only tinted window incidents and using an indicator for multiple tinted window citations as the outcome variable.[34]

[34] Again, given that the binary dependent variable, I use a Probit specification for these regressions. Results are again essentially the same if an OLS specification is used.

Table 14 - Difference Between Actual and Expected Likelihood of Multiple Tinted Window Citations within Tinted Window Incidents					
	Coeff. On Predictors of Multiple Tinted Window Citations for Non-minority Drivers (Probit)				
Variable	Pre-Strikeforce Era	Strikeforce Era	BTVA Era	Post-Strikeforce Era	Current Era
Age	0.012	-0.001	-0.005	0.011	-0.018
	(0.014)	(0.004)	(0.003)	(0.009)	(0.018)
12PM-6PM	3.719***	0.247	0.946***	1.272***	1.292**
	(0.353)	(0.198)	(0.123)	(0.382)	(0.505)
6PM-12AM	4.296***	-0.049	0.894***	1.397***	1.387***
	(0.167)	(0.220)	(0.146)	(0.394)	(0.533)
12AM-6AM	4.122***	-0.085	-0.269	0.639	1.768**
	(0.584)	(0.295)	(0.348)	(0.542)	(0.717)
Constant	-5.334***	0.018	0.118	-1.261***	-0.970*
	(0.367)	(0.198)	(0.143)	(0.450)	(0.529)
Observations	105	1,396	924	275	112
	Mean Difference Between Actual and Predicted Likelihood of Multiple Tinted Window Citations w/in Tinted Window Incidents				
Minority Drivers	0.071	0.124***	0.095***	0.106***	-0.011
	(0.044)	(0.026)	(0.017)	(0.035)	(0.039)
Driver Race Unrecorded	0.032	0.070***	0.042***	0.120***	-0.105*
	(0.057)	(0.025)	(0.016)	(0.037)	(0.058)
Intercept (Non-minority Drivers)	-0.000	-0.000	-0.000	-0.000	-0.001
	(0.036)	(0.032)	(0.019)	(0.034)	(0.038)
*"Excluded" Comparison Time Category 6AM-12PM. All standard errors are heteroskedasticity robust and clustered by census tract. ***indicates significance at 1% level, **indicates significance at 5% level, *indicates significance at 10% level.					

299. Once again, the top panel of Table 14 shows the relationship between driver age and time-of-day and the likelihood of multiple tinted window tickets in tinted window incidents for those recorded as Non-Minorities. As can be seen, these relationships vary somewhat across eras. The most consistent predictor of multiple tinted windows tickets in tinted window incidents among Non-Minorities is whether the incident occurred between 12PM and 6PM. Looking at the bottom panel of Table 14, we can see that conditional on time of day and driver age, during the Strikeforce, BTVA, and Post-Strikeforce eras, the actual likelihood of receiving multiple tinted window citations in tinted window incidents is 10 to 12 percentage points higher for those classified as Minorities than would be predicted if such individuals had been treated the same as Non-Minorities. These differences are all statistically significant at the 1% level. Even in the Pre-Strikeforce and Current eras, these coefficients are positive, though not statistically significant at standard levels. Results for drivers of unrecorded race are generally about half the magnitude of those for Minority drivers, which is consistent with the estimate that drivers of unrecorded race include a roughly equal mix of Minority and Non-Minority drivers.

C. Analysis of Racial Disparities in Stops Data

300. Starting in 2020, the BPD agreed to document traffic stops that led to a stop receipt but no actual citation. The data do not indicate why these stops did not lead to citation tickets. While the BPD could record the race of driver associated with these stop receipts, such information is missing in just under 20% of observations. Moreover, the stop receipts data lacks information about the driver's zip code of residence for the vast majority of observations, so I cannot do a procedure analogous to what I did with respect to citations to impute the racial composition of those for whom race is unrecorded. Although I am therefore limited in the data that I can consider, it is still informative to examine these data.

301. My results show that the reasons given for such stops differ between the BPD Districts with higher Minority populations relative to those with lower Minority populations. Stops in Districts with higher Minority populations were more likely to be for equipment violations or for failure to use a turn signal, while stops in Districts with lower Minority populations were more likely to be for failure to stop at stop signs or traffic lights, or for erratic driving. Similar discrepancies with respect to the reasons why drivers were stopped also exist based on the race of the driver and not just on where these stops happened. These results are described more specifically below.

302. The first two columns of numbers in Table 15 show the number of stop receipts issued, and fraction of total receipts issued, by BPD District. As can be seen, just over 50% of these stop receipts were issued in BPD Districts C and E (which cover eastern and northeastern Buffalo). As can be seen in the two right-most columns of Table 15, these two districts cover just 37% of the total population of Buffalo, but are the two highest Minority districts, at 59% and 73% respectively. Indeed, if we just look at Districts A and E, the Districts with the lowest and highest fraction of Minority residents respectively, we can see that over three times as many stop receipts were issued in the latter than the former.

Table 15 - Stop Receipts by District

District	Num of Stops	% of stops	% Minority	% of City Pop
A	872	8%	17%	16%
D	2,499	23%	31%	29%
B	1,745	16%	50%	17%
C	2,717	25%	59%	14%
E	2,953	27%	73%	23%

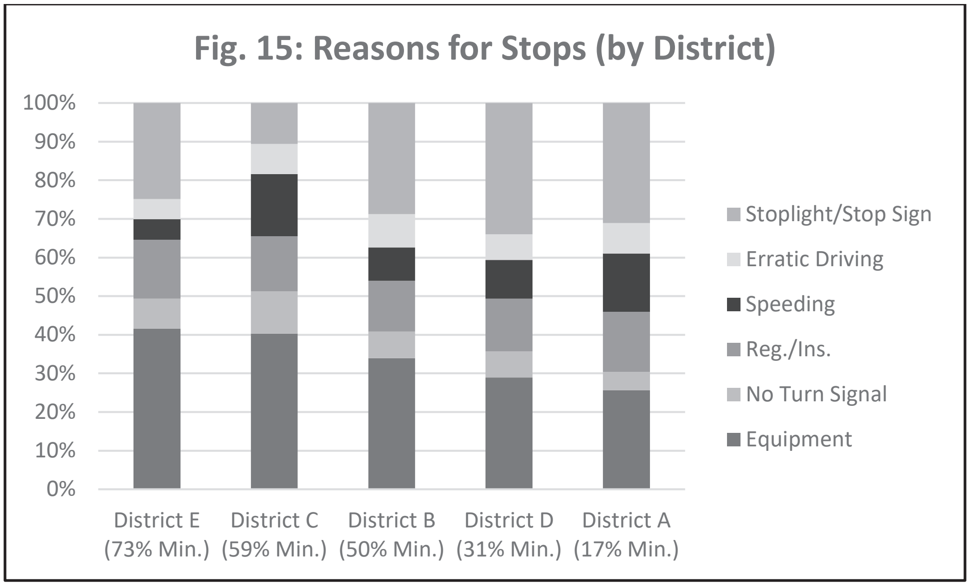
303. As stated above, race of driver is not recorded for just under 20% of the population. However, among stops for which race of driver is recorded, 73% of the stopped individuals were Minority. This coincides with the TraCS data on citations analyzed earlier, in which 73% of stops that led to a citation during a similar time period also involved the stop of a Minority individual. Thus, given that Minorities make up slightly less than half of the Buffalo population, in the Current era, Minority residents of Buffalo were stopped and issued citations, or stopped and issued a stop receipt, more than twice as often as Non-Minority residents of Buffalo, relative to their populations.

304. As noted, the stops data lists the reason for the stop by the BPD officer who conducted the stop. Table 16 documents the reason given for the stops. As can be seen, 94% of the stops were described as being conducted for equipment issues, failure to stop, failure to signal, registration or insurance violations, speeding, or erratic/reckless/illegal driving. Among these top six categories, race of driver was unrecorded about 20% of the time in each category.

Table 16 - Summary of Stop Receipts

Reason For Stop	Number of Stops	Frac. Missing Race
Equipment	3,574	0.19
Stoplight/Stop Sign Violation	2,447	0.20
Registration/Insurance	1,430	0.19
Speeding	1,037	0.19
Failure to use Turn Signal	793	0.21
Erratic/Reckless/Ilegal driving	690	0.19
Investigation/Stolen/Warrant	141	0.09
Suspicion/Suspected/Drugs	101	0.10
Illegal Parking/Stopped	89	0.18
Cell Phone/Texting	65	0.32
Seatbelt/Childseat	43	0.47
No Reason Given	234	0.11
Total	10,644	0.19

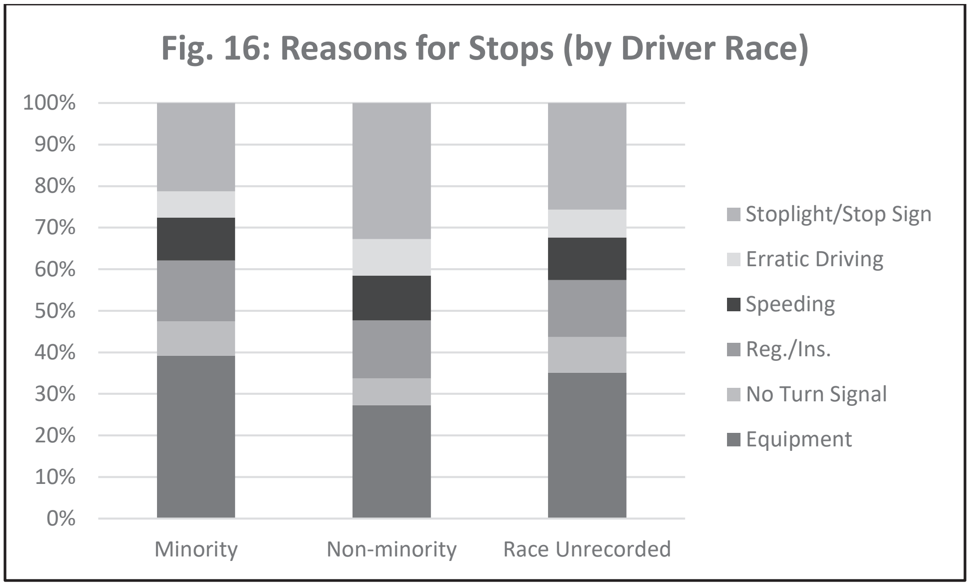
305. Figure 15 considers just stops under these top six reasons for stop. Together, these six reasons for stops account for 94 % of all stops, and there is a large drop-off between the sixth most common reason for stops (690 stops) and the seventh (141 stops). Figure 15 shows the fraction of stops under each of these six categories by BPD District. The Districts are ordered with the highest Minority District on the far left, moving to the lowest Minority District on the far right.



306. As can be seen, the composition of stops by reason for stop differs across districts. Notably, in the two Districts with the largest Minority population in percentage terms (District E and District C), around 50% of stop receipts were given for equipment issues or failure to signal. By contrast, in the two Districts with the lowest Minority population in percentage terms (District D and District A), less than 35% of stop receipts were given for equipment issues or failure to signal.

307. One the other hand, erratic driving or failure to stop at a stoplight or stop sign accounted for 30% and 20% of stops, respectively, in District E and District C, the two highest-Minority Districts. This compares to around 40% of stop receipts given for erratic driving or failure to stop at a stoplight or stop sign in the two lowest-Minority Districts (D and A).

308. Although race of driver data is missing for around 20% of observations, a similar picture arises when looking by race for those observations for which race is recorded. Notably, Figure 16 shows the composition of stops by reason for stop, among stops for the top six reasons, by race.



309. Comparing the first column, which corresponds to drivers recorded as Minorities, to the second column, which corresponds to drivers recorded as Non-Minorities, shows that about 50% of stop receipts issued to those recorded as Minorities being for equipment issues or failure to use turn signal. This compares to less than 35% of stop receipts issued to those recorded as Non-Minorities for these two reasons. Similar to what was shown in Figure 15, Figure 16 shows that although less than 30% of stop receipts issued to those recorded as Minorities were for erratic driving or failure to stop at a stoplight or stop sign, over 40% of stop receipts issued to those recorded as Non-Minorities were issued for these two reasons.

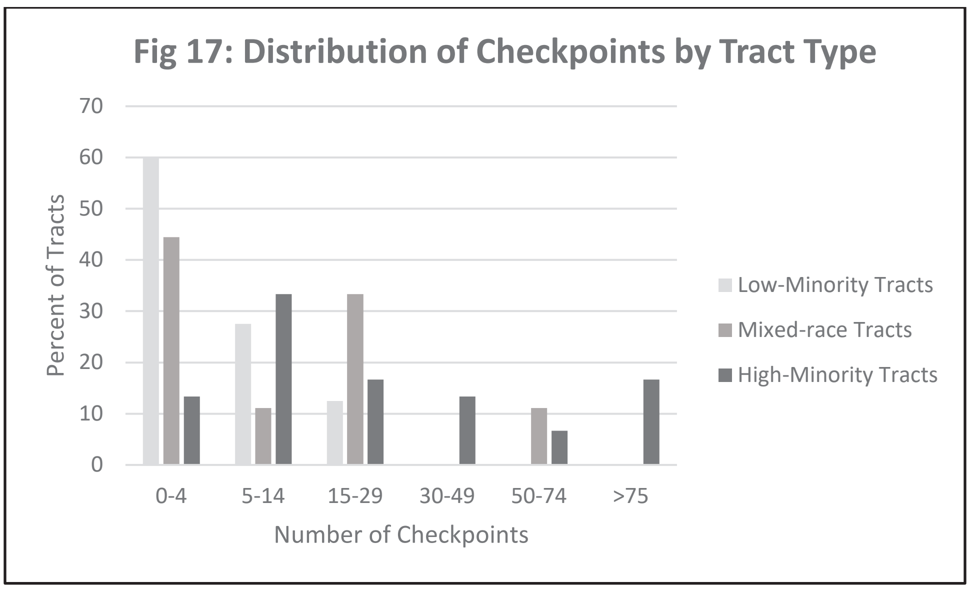
310. Finally, the third column in Figure 16 shows the distribution of reasons for stop for those stops with unrecorded driver race. As can be seen, the fraction of stops for equipment or failure to signal is smaller among those of unrecorded race than it is among those who were recorded as Minorities, but larger than it is among those who were recorded as Non-Minorities. Similarly, the fraction of stops for erratic driving or failure to stop at a stoplight or stop sign among those of unrecorded race was larger than it is among those recorded as Minorities, but smaller than it is among those recorded as Non-Minorities. This is consistent with those of unrecorded race constituting a mix of Minorities and Non-Minorities.

D. Analysis of Racial Bias in Checkpoint Locations

311. Between January 1, 2013, and October 27, 2017, the BPD set up over 1,600 checkpoints. This section analyzes the location of these checkpoints in more detail, particularly the relationship between the location of these checkpoints and the racial demographics of

neighborhoods surrounding such locations. My results show that the BPD overwhelmingly located these checkpoints in higher Minority neighborhoods. Moreover, this greater tendency to setup checkpoints in higher Minority neighborhoods cannot be fully explained by the BPD simply choosing to locate checkpoints in neighborhoods with high crime or a high number of accidents.

312. Figure 17 shows the distribution of checkpoints per tract for tracts classified as High-Minority, Mixed-Race, and Low-Minority tracts. This figure reveals a number of things quite clearly. First, 60% of Low-Minority tracts experienced fewer than five checkpoints, less than 15% experienced more than fifteen checkpoints, and none experienced more than thirty checkpoints. By contrast, less than 15% of High-Minority tracts experienced fewer than five checkpoints, and almost 15% experienced more than thirty. Indeed, four High-Minority tracts experienced over 100 checkpoints each. This is particularly notable again given that High-Minority and Low-Minority tracts on average have very similar populations.



313. As discussed in the Methods section above, the first method I use to analyze checkpoint locations by racial composition of the neighborhood is to again focus on just the High Violent Crime (HVC) tracts, to examine the possibility that the BPD may have primarily chosen to locate checkpoints in the highest crime areas. To identify those tracts, I used the tracts that were in the top third of the violent crime distribution in the 12 months leading up to the implementation of the Strikeforce, with checkpoints arising a few months later.

314. As discussed earlier, Table 3 showed that among HVC tracts, High-Minority tracts were quite similar to Low-Minority tracts in terms of Pre-Strikeforce violent crime rates. However,

18 of the 30 High-Minority tracts were in this HVC group, but only 5 of the 40 Low-Minority tracts were in this HVC group (4 of the 9 Mixed-Race tracts were in this group).

315. Table 17 shows the results of comparing the mean number of checkpoints in High-Minority, Mixed-Race, and Low-Minority tracts. The top rows show the results for all tracts, while the second set of rows include just those tracts classified as HVC tracts as defined above. Looking first at High-Minority tracts, it is true that, on average, a much larger number of checkpoints were set up in the High-Minority HVC tracts than High-Minority tracts overall. This was also the case among the Mixed-Race tracts. However, this is not the case with respect to Low-Minority tracts. Looking at just the HVC tracts in the second set of rows, we see that on average just over 60 checkpoints were set up in each of the High-Minority HVC tracts, but on average only 8 were set up in each Low-Minority HVC tract. This is only slightly more than the average of 6 checkpoints per Low-Minority tract overall. This indicates that to the extent the BPD was considering crime (or deterrence or a related desire to be "visible") when deciding where to locate checkpoints, they only did so with respect to deciding where to locate checkpoints within the higher Minority areas of the city.

Table 17 - Mean Checkpoints per Tract, by Tract Type

	High-minority (1)	Mixed-race (2)	Low-minority (3)	Differences (1)-(3)	(2)-(3)
All Tracts	42.5	16.1	6.3	36.2***	9.7
	(10.37)	(7.08)	(1.05)	(10.42)	(7.16)
	n = 30	n = 9	n = 40		
High-Vio. Crime Tracts Only	61.0	29.4	8.3	52.7***	21.1
	(15.78)	(12.33)	(3.26)	(16.11)	(12.75)
	n = 18	n = 4	n = 5		
High-Vio. Crime Tracts Only (Jul 2012-Jun 2017)	57.7	20.6	3.1	54.6***	17.5*
	(14.99)	(9.67)	(1.01)	(15.03)	(9.73)
	n = 18	n = 4	n = 5		
High-Vio. Crime Tracts Only (Jul 2017 - Dec 2017)	3.3	8.8	5.2	-1.9	3.6
	(1.01)	(2.84)	(2.46)	(2.66)	(3.75)
	n = 18	n = 4	n = 5		
High-Vio. Crime Tracts Only (excluding unexecuted checkpoints)	50.0	22.5	6.6	43.4***	15.9
	(13.24)	(11.28)	(2.56)	(13.48)	(11.57)
	n = 18	n = 4	n = 5		

Note: High-Violent Crime tracts are tracts that were in the top third of the violent crime distribution in 12-mo. Leading up to implementation of Strikeforce. There were 373 reported checkpoints in which no tickets appear to have been issued. I refer to these as "unexecuted" checkpoints.

316. The fourth column of numbers in Table 17 shows that the difference between High-Minority and Low-Minority tracts in terms of average number of checkpoints per tract, and the fifth column shows the difference between Mixed-Race tracts and Low-Minority tracts. As can be seen in the fourth column, this difference is statistically significant at the 1% level between all High-Minority and Low-Minority tracts, and between HVC High-Minority and HVC Low-Minority tracts. Indeed, the average number of checkpoints in each High-Minority HVC tract is over six times higher than the average number of checkpoints per tract in each Low-Minority HVC tract. This again indicates that whether or not a tract was High-Minority is a much stronger predictor of checkpoints than whether or not a tract had high rates of violent crime.

317. I also consider the possibility that BPD behavior with respect to checkpoint location choice may have changed after the Buffalo Common Council adopted a resolution in July 2017

raising concerns about BDP's potentially discriminatory use of checkpoints. To consider that possibility, the third set of rows of Table 17 focuses on just checkpoints implemented prior to July 2017, when this resolution was announced. As can be seen, just focusing on this time period, exceedingly few checkpoints were set up in Low-Minority tracts, even Low-Minority tracts with relatively high rates of violent crime. Rather, almost all checkpoints were located in High-Minority or Mixed-Race tracts during this time frame. By contrast, the fourth set of rows on Table 17 focus on checkpoints implemented after this resolution was announced. As can be seen, while there were not very many checkpoints implemented during this time period in all tract categories, there were actually more checkpoints implemented in HVC Mixed-Race and HVC Low-Minority tracts than HVC High-Minority after the resolution.

318. While my analysis of checkpoints so far has focused on all the checkpoints documented by the BPD, there are also 373 ostensible checkpoints that do not appear to correspond to the issuance of actual citations in the TraCS data. I express no opinion about whether it is appropriate to include or exclude those checkpoints from the analysis of whether BPD's decisions about where to locate checkpoints were indicative of racial basis. However, for completeness, I have also analyzed the checkpoints data excluding these checkpoints in which no tickets were issued. The bottom rows of numbers in Table 17 show what happens when I exclude these checkpoints. As can be seen, the results are largely similar, showing that BPD conducted far more checkpoints that resulted in tickets per tract in High-Minority HVC tracts than in Low-Minority HVC tracts.

319. Overall, Table 17 shows that even if we examine only those tracts with high rates of violent crime, High-Minority tracts experienced over seven times more checkpoints on average than Low-Minority tracts with similar violent crime rates.

320. My second approach to assessing racial disparities in checkpoint location is the regression prediction approach, where I estimate how accidents and criminal activity relate to checkpoint counts in Low-Minority tracts, and use these estimates to predict the number of checkpoint counts that should arise in High-Minority (and Mixed-Race) tracts if the relationship between prior accidents and criminal activity and subsequent checkpoint deployment was the same in High-Minority (and Mixed-Race) tracts as it was in Low-Minority tracts. I then calculate the difference between actual and predicted checkpoints counts for each tract to see how much these differ on average.

321. Table 18 shows the results of this exercise. The first column of the top panel shows how population, prior accidents, and criminal incidents correlate with checkpoints in Low-Minority tracts over the entire checkpoint period. As can be seen, the strongest predictors of checkpoints in Low-Minority tracts are population size of tract and number of prior violent criminal incidents.

Table 18 - Diff. Between Actual and Predicted Checkpoints per Tract

	Predictors of Num. of Checkpoints in Low-minority Tracts (OLS)			
Variable	**Jul 2012-Dec 2017**	**Jul 2012-Jun 2017**	**Jul 2017 - Dec 2017**	**Excl. Unexecuted Chkpts**
Total Population (100s)	0.220**	0.074	0.146***	0.125*
	(0.096)	(0.095)	(0.040)	(0.066)
minor accidents (10s)	-0.855	-0.671	-0.184	-0.486
	(0.721)	(0.690)	(0.365)	(0.434)
injury accidents (10s)	0.866	0.715	0.151	0.338
	(0.930)	(0.757)	(0.512)	(0.659)
violent incidents (100s)	4.962**	2.510	2.453**	3.795***
	(2.031)	(1.760)	(1.087)	(1.316)
property incidents (100s)	-1.816**	-0.639	-1.177**	-1.204*
	(0.786)	(0.624)	(0.550)	(0.594)
Constant	-0.854	0.488	-1.342	0.026
	(3.235)	(3.031)	(1.497)	(2.174)
Observations	40	40	40	40
R-squared	0.284	0.132	0.299	0.265
	Avg. Diff. Between Actual and Predicted Checkpoints per Tract			
High-minority Tracts (N = 30)	33.154***	35.116***	-1.962**	27.581***
	(10.314)	(9.882)	(0.860)	(8.677)
Mixed-race Tracts (N = 9)	8.000	6.498	1.503	6.716
	(6.383)	(4.979)	(1.857)	(5.761)
Intercept (Low-minority Tracts)	0	-0.000	0.000	0.000
	(0.897)	(0.655)	(0.499)	(0.625)

Heteroskedastic robust standard errors shown in parentheses. Accidents and criminal incidents correspond to the numbers of such incidents in the year prior to the implementation of the Strikeforce. ***indicates significance at 1% level, **indicates significance at 5% level, *indicates significance at 10% level. There were 373 reported checkpoints in which no tickets appear to have been issued. I refer to these as "unexecuted" checkpoints.

322. Moving to the bottom panel of Table 18, the first column of the top row shows that, on average, between July 2012 and December 2017, High-Minority tracts experienced over 33 more checkpoints each than would be predicted by how such checkpoints were allocated in Low-Minority tracts. This result is significant at the 1% level.

323. The second and third columns of Table 18 show the results of the exercise described above but separating the checkpoints that occurred before this Buffalo Common Council action in July 2017 and after. As can be seen in the second column in the top panel, in this time periods prior to July 2017, prior violent crime rates are not correlated with checkpoint location among

Low-Minority tracts, but they are after. Moreover, the R-squared (a measure of how much of the variance in checkpoints across Low-Minority tracts is explained by the predicting variables) increases from 0.132 to 0.299 between the pre-July 2017 and post-July 2017 samples, implying that it was only after the Common Council resolution that prior violent crime appear to be correlated with where checkpoints were located in lower Minority neighborhoods.

324. Looking at the second column of the bottom panel shows an even greater difference between the actual and predicted number of checkpoints in High-Minority tracts in the pre-July 2017 sample than in the overall sample, with the third column of the lower panel showing the opposite to be true in the post-July 2017 sample. This again indicates that, prior to the July 2017 Buffalo Common Council action, the primary predictor of where checkpoints were located was the fraction of a neighborhood that was Minority, not previous crime rates.

325. The final column of Table 18 shows what happens when I exclude those checkpoints in which no tickets appear to have been issued (i.e., "unexecuted checkpoints") from the analysis over the July 2012 to December 2017 period. Again, my results change only slightly when I exclude these checkpoints.

326. As with the analysis of citations by census tract category, one might be worried that the exercise above places undue reliance on the linearity assumptions embedded in OLS when predicting checkpoints for High-Minority tracts based on how prior accidents and crime correlated with checkpoints in Low-Minority tracts. In other words, one might question whether one should assume that checkpoint frequency scales up linearly as crime and accident rates rise. Therefore, as a final complementary analysis, I also present the results of the Propensity Score matching approach discussed in the Methods section. Before presenting results from this analysis, recall that one of the constraints to this method is that we can only include tracts that are in the "common" support, or in other words we must exclude High-Minority tracts that are very unlike any Low-Minority tracts in terms of their crime rates and accident rates. As discussed previously, this excludes 12 of the 30 High-Minority tracts, where these excluded tracts generally had higher rates of property and violent crime prior to the implementation of the Strikeforce.

327. Table 19 presents the results of the propensity score exercise. The top panel presents the results for the full checkpoint time-period. The top row ("Unmatched") again shows how the average number of checkpoints differed between High-Minority and Low-Minority tracts overall. The second row ("Matched") shows how the average number of checkpoints differed

between High-Minority tracts and matched Low-Minority tracts based on tract population, prior accidents and criminal incidents, again using only the subset of High-Minority tracts that lied within the common support. As can be seen, the difference in average checkpoints between High-Minority and Low-Minority tracts is smaller among the matched tracts than overall, which is not surprising given that this analysis excludes the High-Minority tracts that cannot be matched to any Low-Minority tracts, which included several of the tracts that experienced some of the highest numbers of checkpoints. Still, High-Minority tracts in the common support experienced an average of almost 17 more checkpoints per tract than Low-Minority tracts with similar numbers of prior accidents and criminal incidents. However, due to relatively large standard errors, the p-value on this estimated difference is 0.16, which slightly exceeds standard levels of significance.

Table 19 - Propensity Score Matching Estimates of Checkpoints by Tract type

	Avg. Checkpoints per Tract					
Sample	High Minority	Low Minority	Diff.	S.E.	T-stat	p-val
Full Checkpoint Period (Jul 2012-Dec 2017)						
Unmatched	42.5	6.3	36.2	10.5	3.45	0.00
Matched	25.7	9.9	15.7	11.0	1.43	0.164
Pre-Comm. Council Resolution (Jul 2012-Jun 2017)						
Unmatched	40.2	3.1	37.1	10.0	3.7	0.00
Matched	23.9	3.6	20.3	10.4	1.95	0.061
Post-Comm. Council Resolution (Jul 2017-Dec 2017)						
Unmatched	2.4	3.3	-0.9	0.9	-1.0	0.32
Matched	1.8	6.4	-4.6	1.9	-2.48	0.020
Excluding Unexecuted Checkpoints (Jul 2012-Dec 2017)						
Unmatched	34.2	4.4	29.8	8.8	3.4	0.00
Matched	20.0	7.2	12.8	9.2	1.39	0.175

Note: Standard errors are calculated to be robust to heteroskedasticity. There were 373 reported checkpoints in which no tickets appear to have been issued. I refer to these as "unexecuted" checkpoints.

328. The second panel down on Table 19 shows that these racial discrepancies in checkpoints across tracts are even larger (especially in percentage terms) if we limit the sample to checkpoints occurring before the Buffalo Common Council resolution in July 2017. When looking only at these pre-July 2017 checkpoints, High-Minority tracts in the common support experienced significantly more checkpoints than matched Low-Minority tracts at below the 5% level.

329. The third panel down on Table 19 shows the results for the time after the Buffalo Common Council Resolution. Consistent with the previous results, we see that during this timeframe, there were relatively few checkpoints implemented and High-Minority tracts in the

common support actually experienced significantly fewer checkpoints than matched Low-Minority tracts at below the 5% level.

330. The bottom panel of Table 19 shows what happens when I exclude the unexecuted checkpoints in which no tickets appear to have been issued. Again, the results here are quite similar to those shown in the top panel which included all checkpoints.

331. In summary, this subsection shows that up until the July 2017 Buffalo Common Council Resolution that raised concerns about racial discrimination with respect to the use of checkpoints, the BPD located far more checkpoints in neighborhoods with a high Minority make-up than neighborhoods with low levels of Minorities. Moreover, these discrepancies cannot be fully explained by differences in the rates at which criminal incidents or traffic accidents occurred between High-Minority neighborhoods and Low-Minority neighborhoods.

Respectfully submitted this 29th day of May,

David Bjerk, Ph.D

APPENDIX: ADDITIONAL ANALYSES

Table A1a - Difference Between Actual and Expected Total Citations per Month (Pre-Strikeforce Controls)					
	Coefficients on Predictors of Total Monthly Citations in Low-minority Tracts (OLS)				
Variable	**Pre-Strikeforce Era**	**Strikeforce Era**	**BTVA Era**	**Post-Strikeforce Era**	**Current Era**
Total Population (100s)	0.06	0.05	-0.12	-0.34*	-0.23*
	(0.10)	(0.11)	(0.20)	(0.20)	(0.11)
Pre-Strikeforce minor accidents (10s)	-1.44	-2.09**	-1.35	0.63	0.03
	(0.89)	(0.91)	(1.92)	(1.92)	(1.14)
Pre-Strikeforce injury accidents (10s)	2.43	3.89***	2.20	2.62	1.39
	(1.48)	(1.25)	(2.06)	(3.43)	(1.85)
Pre-Strikeforce vio. incidents (100s)	6.54**	10.50***	20.47***	11.76**	11.07**
	(2.59)	(2.66)	(5.43)	(5.78)	(4.14)
Pre-Strikeforce prop. incidents (100s)	1.17	1.93	0.53	1.33	1.18
	(1.42)	(1.30)	(2.09)	(3.09)	(1.58)
Constant	4.95	2.63	12.28*	15.96	6.78
	(3.78)	(3.77)	(7.03)	(9.55)	(4.56)
Observations	240	1,440	1,240	1,000	1,360
R-squared	0.24	0.33	0.32	0.37	0.37
	Avg. Difference Between Actual and Predicted Monthly Citations by Tract Type				
High-minority Tracts (N = 30)	-6.82***	29.79***	51.61***	6.62	-6.28**
	(2.04)	(9.62)	(16.19)	(5.59)	(2.42)
Mixed-race Tracts (N = 9)	2.27	11.74*	13.08	2.14	1.85
	(5.09)	(6.97)	(12.46)	(5.59)	(4.38)
Intercept (Low-Minority Tracts)	0.00	0.00	-0.00	0.00	-0.00
	(1.30)	(1.19)	(2.19)	(2.61)	(1.42)

Heteroskedastic robust standard errors clustered by census tract shown in parentheses. Accidents and criminal incidents correspond to the numbers of such incidents in the year prior to the start of Strikeforce. ***indicates significance at 1% level, **indicates significance at 5% level, *indicates significance at 10% level.

Table A1b - Difference Between Actual and Predicted Monthly Non-moving Citations					
	Coefficients on Predictors of Non-Moving Cit. per 6-mo. in Low-minority Tracts (OLS)				
Variable	**Pre-Strikeforce Era**	**Strikeforce Era**	**BTVA Era**	**Post-Strikeforce Era**	**Current Era**
Total Population (100s)	-0.01	0.01	-0.02	-0.20*	-0.16**
	(0.07)	(0.10)	(0.18)	(0.10)	(0.06)
Pre-Strikeforce minor accidents (10s)	-1.28*	-1.70**	-1.30	0.11	-0.19
	(0.67)	(0.79)	(1.68)	(0.99)	(0.71)
Pre-Strikeforce injury accidents (10s)	1.41	2.55**	1.87	0.82	0.79
	(1.04)	(1.05)	(1.75)	(1.50)	(0.96)
Pre-Strikeforce vio. incidents (100s)	6.94***	9.72***	16.24***	10.72***	8.71***
	(1.79)	(2.26)	(4.84)	(3.32)	(2.68)
Pre-Strikeforce prop. incidents (100s)	-0.04	0.77	0.61	0.29	0.86
	(0.96)	(1.01)	(1.83)	(1.41)	(0.81)
Constant	4.73*	2.93	5.87	8.09*	2.41
	(2.74)	(3.45)	(5.68)	(4.23)	(2.33)
Observations	240	1,440	1,240	1,000	1,360
R-squared	0.27	0.29	0.31	0.42	0.41
	Avg. Diff. Between Actual and Predicted Monthly Non-moving Citations				
High-minority Tracts (N = 30)	-4.36***	27.85***	47.37***	9.09**	-3.18*
	(1.40)	(8.15)	(14.23)	(4.28)	(1.76)
Mixed-race Tracts (N = 9)	0.86	9.49*	13.27	3.35	1.69
	(2.88)	(5.26)	(11.06)	(3.70)	(2.51)
Intercept (Low-Minority Tracts)	0.00	0.00	-0.00	0.00	0.00
	(0.89)	(1.00)	(1.85)	(1.35)	(0.78)

Heteroskedastic robust standard errors clustered by census tract shown in parentheses. Accidents and criminal incidents correspond to the numbers of such incidents in the year prior to the start of Strikeforce. ***indicates significance at 1% level, **indicates significance at 5% level, *indicates significance at 10% level. Non-moving violation citations include citations for equipment infractions (e.g., tinted windows, obscured view, seatblelt/childseat infractions, worn tires, broken lights), license/registration infractions, insurance infractions, and inspection infractions.

Table A1c - Difference Between Actual and Predicted Monthly Moving Citations					
	Coeff. on Predictors of Monthly Moving Citations in Low-minority Tracts (OLS)				
Variable	**Pre-Strikeforce Era**	**Strikeforce Era**	**BTVA Era**	**Post-Strikeforce Era**	**Current Era**
Total Population (100s)	0.07	0.05	-0.08	-0.05	-0.04
	(0.07)	(0.03)	(0.05)	(0.10)	(0.07)
Pre-Strikeforce minor accidents (10s)	-0.17	-0.39*	-0.28	0.04	0.00
	(0.44)	(0.23)	(0.32)	(0.98)	(0.50)
Pre-Strikeforce injury accidents (10s)	1.04	1.34***	0.65	2.22	0.72
	(0.87)	(0.36)	(0.53)	(1.81)	(0.84)
Pre-Strikeforce vio. incidents (100s)	-0.45	0.70	4.30***	1.91	2.27
	(1.52)	(0.89)	(1.02)	(2.80)	(1.75)
Pre-Strikeforce prop. incidents (100s)	1.25	1.22***	0.21	1.25	0.72
	(0.86)	(0.43)	(0.58)	(2.22)	(0.81)
Constant	0.12	-0.45	4.83**	2.52	2.13
	(2.34)	(0.95)	(1.92)	(4.15)	(2.67)
Observations	240	1,440	1,240	1,000	1,360
R-squared	0.13	0.27	0.15	0.23	0.18
	Avg. Diff. Between Actual and Predicted Monthly Moving Citations				
High-minority Tracts (N = 30)	-2.41**	2.09	4.29**	-1.49	-2.36**
	(0.94)	(1.63)	(2.07)	(1.58)	(0.95)
Mixed-race Tracts (N = 9)	1.52	2.42	0.22	-0.16	0.13
	(2.77)	(1.79)	(1.93)	(2.13)	(1.54)
Intercept (Low-Minority Tracts)	0.00	0.00	-0.00	0.00	0.00
	(0.68)	(0.35)	(0.52)	(1.15)	(0.72)

Heteroskedastic robust standard errors clustered by census tract shown in parentheses. Accidents and criminal incidents correspond to the numbers of such incidents in the year prior to the start of Strikeforce. ***indicates significance at 1% level, **indicates significance at 5% level, *indicates significance at 10% level. Moving violation citations include moving violations (e.g., speeding, reckless driving, failure to stop at red light or stop sign), DUIs, and mobile electronics infractions.

EXHIBIT A: CURRICULUM VITAE (MAY 2024)

DAVID J. BJERK

Robert Day School of Economics and Finance
Claremont McKenna College
Claremont, CA 91711-6400
Email: david.bjerk @ cmc. edu
Website: https://sites.google.com/site/davidbjerk/home

Employment

Claremont McKenna College – Robert Day School of Economics and Finance

Department Chair, July 2022 – present.

Russell S. Bock Professor of Public Economics, 2019 – present.

Associate Professor of Economics, 2011 – 2018.

Assistant Professor of Economics, July 2007 – 2010.

RAND Corporation
Research Fellow in Population Studies, 2005 - 2007.

McMaster University
Assistant Professor of Economics, 2003 –2007.

Other Affiliations

Fellow of the Society for Empirical Legal Studies, 2020 – present.

Research Fellow, IZA (Institute for Study of Labor), 2009 – present.

Education

Ph.D., Economics, University of Wisconsin-Madison.

M.S., Economics, University of Wisconsin-Madison.

B.A., Economics, Carleton College.

Research Areas of Interest

Criminal Justice, Labor Economics, Public Economics, Urban Poverty.

Scholarly Publications

Bjerk, David and Shawn Bushway. (2023). "The Long-term Incarceration Consequences of Growing Up in a Crime Boom." *Journal of Quantitative Criminology* 39: 1003-1025.

Bjerk, David. (2022). "Does Greater Police Funding Help Catch More Murderers?" *Journal of Empirical Legal Studies* 19(3): 528-559. (Lead Article)

Bjerk, David. (2021). "Socially Optimal Plea Bargaining with Costly Trials and Bayesian Juries." *Economic Inquiry* 59(1): 263-279.

Bjerk, David and Eric Helland. (2020). "What Can DNA Exonerations Tell Us About Racial Differences in Wrongful Conviction Rates?" *Journal of Law and Economics* 63(2): 341-366.

Bjerk, David. (2019). "Replication of Mismatch Research: Ayers, Brooks, and Ho (comment)." *International Review of Law and Economics* 58: 3-5.

Bjerk, David and I. Serkan Ozbeklik. (2018). "Using Samples-of-Opportunity to Assess Gender Bias in Principal Evaluations of Teachers: A Cautionary Tale." *Journal of Labor Research* 39(3): 235-258. (Lead Article)

Bjerk, David. (2017). "Mandatory Minimums and the Sentencing of Federal Drug Crimes." *Journal of Legal Studies* 46(1): 93-128.

Bjerk, David. (2017). "Mandatory Minimum Reform and the Sentencing of Crack Cocaine Defendants: An Analysis of the Fair Sentencing Act." *Journal of Empirical Legal Studies* 14(2): 370-396.

Bjerk, David. (2016). "In Front of and Behind the Veil of Ignorance: An Analysis of Motivations for Redistribution." *Social Choice and Welfare* 47(4): 791-824.

Bjerk, David and Caleb Mason. (2014). "The Market for Mules: Risk and Compensation of Cross Border Drug Couriers" *International Review of Law and Economics* 39: 58-72.

Bjerk, David and Caleb Mason. (2013). "Inter-Judge Sentencing Disparity on the Federal Bench: An Examination of Drug Smuggling Cases in the Southern District of California." *Federal Sentencing Reporter* 25(3): 190-196.

Bjerk, David. (2012). "Re-examining the Impact of Dropping Out on Criminal and Labor Market Outcomes in Early Adulthood." *Economics of Education Review* 31(1): 110-122.

Bjerk, David. (2010). "Thieves, Thugs, and Neighborhood Poverty." *Journal of Urban Economics* 68: 231-246.

Bjerk, David. (2009). "How Much Can We Trust Causal Interpretations of Fixed Effects Estimators in the Context of Criminality?" *Journal of Quantitative Criminology* 25(4): 391-417.

Bjerk, David. (2009). "Beauty vs. Earnings: Gender Differences in Human Capital, Earnings, and Priorities Over Spousal Characteristics in a Matching Model." *Journal of Economic Behavior and Organization* 69: 248-259.

Bjerk, David. (2008). "Glass Ceilings or Sticky Floors? Statistical Discrimination in a Dynamic Model of Promotion and Hiring," *Economic Journal* 118(July): 961-982.

Bjerk, David. (2008). "On the Role of Plea Bargaining and the Distribution of Sentences in the Absence of Judicial System Frictions." *International Review of Law and Economics* 28(1): 1-7 (Lead Article).

Bjerk, David. (2007) "Guilt Shall Not Escape or Innocence Suffer? The Limits of Plea Bargaining When Defendant Guilt is Uncertain," *American Law and Economics Review* 9(2):305-329 (Lead Article).

Bjerk, David and Seungjin Han. (2007). "Assortative Marriage and the Effects of Government Homecare Subsidy Programs on Gender Wage and Participation Inequality," *Journal of Public Economics* 91(5-6): 1135-1150.

Bjerk, David. (2007). "Racial Profiling, Statistical Discrimination, and the Effect of a Colorblind Policy on the Crime Rate," *Journal of Public Economic Theory* 9(3): 543-567.

Bjerk, David. (2007). "The Differing Nature of Black-White Wage Inequality Across Occupational Sectors." *Journal of Human Resources* 42(2): 398-434.

Bjerk, David. (2007). "Measuring the Relationship Between Youth Criminal Participation and Household Economic Resources," *Journal of Quantitative Criminology* 23(1): 23-39.

Bjerk, David. (2005). "Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion Under Mandatory Minimum Sentencing." *Journal of Law and Economics* 48(2): 591-627.

Teaching

Economics of Crime and Criminal Justice (Undergraduate), Claremont McKenna College.

Economics of Poverty, Inequality, and Discrimination (Undergraduate), Claremont McKenna College.

Seminar and Tutorial in Economics for Philosophy, Politics, and Economics (PPE) Program (Undergraduate), Claremont McKenna College.

Intermediate Microeconomic Theory (Undergraduate), Claremont McKenna College.

Senior Research Seminar (Undergraduate), Claremont McKenna College.

Public Economics (PhD Level), McMaster University.

Economics of the Labor Market (Undergraduate), McMaster University.

Professional Service

Board of Directors, *Society for Empirical Legal Studies (SELS)*, 2018-2022.

Area Organizer, *American Law and Economics Association (ALEA) - Criminal Law, Crime, Law Enforcement Sessions,* 2021.

External Review Committee, Macalester College, Department of Economics, 2021.

Co-chair, *Conference on Empirical Legal Studies (CELS)*, 2019.

Associate Editor, *International Review of Law and Economics*, September 2018-present.

Area Organizer, *American Law and Economics Association (ALEA) - Criminal Law, Crime, Law Enforcement Sessions,* 2018.

Program Committee Member, *Program in Empirical Legal Studies Replication Conference,* 2018-2019.

External Review Committee, Carleton College, Physical Education, Athletics, and Recreation (PEAR) Department, 2014.

Program Committee Chair - *Southern California Conference in Applied Microeconomics (SoCCAM),* 2014-2015.

Area Organizer, *American Law and Economics Association (ALEA) - Criminal Law, Crime, Law Enforcement Sessions,* 2013.

Referee-
American Economic Review;
American Economic Journal – Applied Economics;
American Economic Journal – Economic Policy;
American Law and Economics Review;
Applied Economics Research Bulletin;
B.E. Journal of Economic Policy and Analysis;
Canadian Public Policy;
Canadian Journal of Economics;
Contemporary Economic Policy;
Economics Bulletin;

Economics of Education Review;
Economic Inquiry;
Economic Journal;
Empirical Economics;
European Economic Review;
Games and Economic Behavior;
International Migration Journal;
International Migration Review;
International Review of Law and Economics;
IZA Journal of Labor Policy;
Journal of Drug Issues;
Journal of Economic Behavior and Organization;
Journal of Empirical Legal Studies;
Journal of Human Resources;
Journal of Human Capital;
Journal of Labor Economics;
Journal of Labor Research;
Journal of Law and Economics;
Journal of Law, Economics, and Organization;
Journal of Legal Studies;
Journal of Policy Analysis and Management;
Journal of Political Economy;
Journal of Population Economics;
Journal of Public Economics;
Journal of Urban Economics;
Journal of Quantitative Criminology;
Justice Quarterly;
Labour Economics;
Law, Probability, and Risk
Social Choice and Welfare;
Southern Economic Journal;
RAND Journal of Economics;
Review of Economics of the Household;
Review of Economics and Statistics;
Review of Economic Studies;
Quarterly Journal of Economics.

2016-2018 *Distinguished Referee* for *Journal of Legal Studies*.

Reviewer- *Oxford University Press; Oxford Bibliographies; National Science Foundation (NSF); Routledge*; *Social Science and Humanities Research Council (SSHRC); Guggenheim Foundation; Israeli Science Foundation (ISF).*

Invited Talks and Presentations

Smith College - Department of Economics, November 2021.

University of California – Irvine, Department of Economics, October 2021.
West Virginia University, Department of Economics, September 2019.
Newkirk Center for Science and Society, UC-Irvine March 2019.
University of California - Riverside, Applied Economics Seminar, March 2019
University of Michigan, Law and Economics Seminar*,* Oct 2018.
University of Southern California, Law and Social Science Seminar, Oct 2018.
University of Waterloo, Waterloo Economics Workshop*,* March 2017.
University of California - Davis, Department of Economics, May 2016.
Claremont Graduate University, Department of Economics and Politics, Nov 2015.
University of Colorado-Denver, Department of Economics, Oct 2015.
University of Houston/Rice University, Departments of Economics, Oct 2013.
Claremont Graduate University, Dept of Economics and Politics, Apr 2013.
University of Quebec at Montreal (UQAM) – Dept. of Economics, December 2011.
University of California – Irvine, Department of Economics, October 2010.
Vanderbilt University, Law School, February 2010.
University of California – Santa Cruz, Department of Economics, January 2010.
San Diego State University, Department of Economics, April 2009.
University of Southern California, Department of Economics, March 2009.
University of California – Berkeley, Goldman School of Public Policy, October 2007.
University of Cambridge (United Kingdom), Department of Economics, June 2007.
Emory University, Department of Economics, April 2007.
RAND Corporation, November 2006.
Claremont McKenna College, Department of Economics, October 2006.
Texas A&M University, Department of Economics, October 2006.
Rice University/University of Houston, Departments of Economics, October 2006.
York University, Department of Economics, September 2006.
University of Quebec at Montreal, Department of Economics, November 2005.
University of British Columbia, Department of Economics, October 2005.
Simon Fraser University, Department of Economics, October 2005.
University of Toronto, Department of Economics, April 2005.
University of Waterloo, Department of Economics, October 2004.
University of Western Ontario, Department of Economics, April 2004.
Brown University, Department of Economics, March 2003.
Washington and Lee University, Department of Economics, February 2003.
McMaster University, Department of Economics, February 2003.
RAND Corporation, January 2003.
University of Kentucky, Department of Economics, January 2003.
University of Wisconsin, Department of Economics, November 2002.

Conference Participation

Institute for Research on Poverty (IRP) Summer Workshop, University of Wisconsin, June 2023 (presenter).
Conference on Empirical Legal Studies, University of Toronto Law School, November 2021 (presenter, discussant).
Association for Public Policy Analysis and Management, Denver, Nov 2019. (presenter)

Program in Empirical Legal Studies Replication Conference, Claremont McKenna College, April 2019 (discussant).
Conference on Empirical Legal Studies, University of Michigan Law School, November 2018 (presenter, discussant).
All-California Labor Economics Conference, University of Southern California, October 2018 (presenter).
American Law and Economics Association Annual Meetings. Boston University Law School, May 2018 (presenter).
Program in Empirical Legal Studies Replication Conference, Claremont McKenna College, April 2018 (discussant).
Southern California Conference in Applied Microeconomics, Claremont McKenna College, March 2018 (discussant).
Conference on Empirical Legal Studies, Duke University Law School, November 2016 (presenter).
Institute for Research on Poverty (IRP) Summer Research Workshop, University of Wisconsin-Madison, June 2016 (presenter).
Southern California Conference in Applied Microeconomics, Claremont McKenna College, April 2016 (presenter).
Institute for Research on Poverty (IRP) Summer Research Workshop, University of Wisconsin-Madison, June 2014 (presenter).
Justice Database Meeting, Center for Policing Equity, Department of Justice: Washington DC, April 2014 (invited participant).
Conference on Empirical Legal Studies, University of Pennsylvania Law School, October 2013 (presenter).
American Law and Economics Association Meetings, Vanderbilt University Law School, May 2013 (presenter).
Southern California Conference in Applied Microeconomics, Claremont McKenna College, April 2013 (discussant).
Conference on Empirical Legal Studies, Northwestern University Law School, November 2011 (presenter)
San Francisco Federal Reserve Bank Applied Micro Conference, July 2010 *(presenter)*
Southern California Applied Economics Workshop, Claremont McKenna College, April 2010 (presenter).
Crime and Population Dynamics Summer Workshop, U. of Maryland, June 2009 (discussant).
Canadian Law and Economics Association, University of Toronto, September 2008 (presenter, session chair).
National Bureau of Economic Research (NBER), Economics of Crime Working Group, July 2008.
Institute for Research on Poverty (IRP) Summer Research Workshop, University of Wisconsin-Madison, June 2008 (presenter).
Crime and Population Dynamics Summer Workshop, U. of Maryland, June 2007 (presenter).
Population Association of America, New York, March 2007 (presenter).
Canadian Law and Economics Association Meetings, University of Toronto, Sept. 2006 (presenter).

Workshop on Criminology and the Economics of Crime, U. of Maryland, June 2006 (presenter).
American Law and Economics Association Meetings, U. California Berkeley, May 2006 (presenter).
Population Association of America, Los Angeles, April 2006 (session chair, discussant).
Canadian Economics Association, McMaster University, May 2005 (session chair).
Canadian Public Economics Research Group, McMaster University, May 2005 (presenter).
American Law and Economics Association Meetings, New York University, May 2005 (presenter).
Meetings of the Southern Economics Association, New Orleans, Nov. 2004 (presenter).
CIBC Project on Human Capital and Productivity, U. of Western Ontario (session chair).
Canadian Public Economics Research Group, University of Toronto, June 2004 (presenter).
Meetings of the Society of Labor Economists, San Antonio, May 2004 (presenter, discussant).
Meetings of the Society of Labor Economists, Toronto, September 2003 (presenter).
American Law and Economics Association Meetings, Harvard University, May 2002 (presenter).

EXHIBIT B: LIST OF MATERIALS CONSIDERED

DATA

- Census Bureau, American Community Survey (neighborhood demographics)
- City of Buffalo
    - CHARMS (stops, crime)
    - Crosswalk of driver's license number proxies to actual driver's license numbers.
    - Open Data (citations, driver demographics)
    - Roadblock Directives (compiled into list of checkpoints)
    - TraCS (citations)
- New York State Department of Transportation (accidents)

OTHER MATERIALS

- Defendants' Amended Responses to Plaintiffs' Fifth Set of Interrogatories and Sixth Request for Production of Documents
- Defendants' Answer to Plaintiffs' Amended Complaint, ECF No. 64
- Mayor Byron Brown, Executive Order 2020-001
- New York Driver's License Suspension Reform Act of 2021
- New York Vehicle & Traffic Laws
- Plaintiffs' Amended Complaint, ECF No. 63
- Training Bulletin, Ent Mobile – Electronic Traffic Stop Receipt, COB560603